VOLUME 37 JANUARY 1951 NUMBER 1

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
OF THE
UNITED STATES OF AMERICA

EDITORIAL BOARD

LINUS PAULING, *Chairman* EDWIN B. WILSON, *Managing Editor*
FRED. E. WRIGHT, *Home Secretary* ROGER ADAMS, *Foreign Secretary*
DOUGLAS WHITAKER, *Chairman of the National Research Council*

J. W. BEAMS
R. E. CLELAND
W. W. RUBEY
HARLOW SHAPLEY
F. E. TERMAN
G. W. BEADLE
E. A. DOISY
G. C. EVANS
C. H. GRAHAM
J. T. PATTERSON
GREGORY BRRIT
L. C. DUNN
C. S. MARVEL
J. VON NEUMANN
W. H. TALIAFERRO

Publication Office: Mack Printing Company, Easton, Pa.
Editorial Office: Harvard School of Public Health, Boston 15
Home Office of the Academy: Washington 25, D. C.

Entered as Second-Class matter, June 21, 1939, at the post-office at Easton, Pennsylvania, under the Act of August 24, 1912. Accepted for mailing at a special rate of postage provided for in section 1103, act of October 3, 1917. Authorized June 15, 1920.

Published Monthly Price per Annum, $5.00

OFFICERS OF THE ACADEMY

PRESIDENT.....D. W. BROWN
VICE PRESIDENT.....EDWIN B. WILSON
FOREIGN SECRETARY.....ROGER ADAMS
HOME SECRETARY.....FRED. E. WRIGHT
TREASURER.....W. J. ROBBINS

INFORMATION TO SUBSCRIBERS

Checks covering subscriptions to the PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES at the rate of $7.50 per annum should be made payable to the National Academy of Sciences, and sent to the Treasurer, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C. Single numbers $0.75. These prices effective January, 1951.

Microfilm copies of the PROCEEDINGS are available to subscribers to the paper edition. The microfilm is a positive, furnished on metal reels, suitably labeled and is distributed at the end of the volume year. Inquiries concerning purchase should be directed to University Microfilms, 313 N. First St., Ann Arbor, Michigan.

Past publications of the NATIONAL ACADEMY OF SCIENCES from 1863 to 1926 and of the NATIONAL RESEARCH COUNCIL from 1916 to 1925 are listed in the Index to the First Ten Volumes of these PROCEEDINGS. This Index appeared as Part 2 of Volume 13, No. 1, January, 1927, and is for sale separately at the price of $1.50. Orders should be sent to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

PAST volumes of the PROCEEDINGS may be obtained at five dollars per volume unless the sale of the volume would break a complete set of Volumes I to XXX. Single numbers may be obtained for fifty cents except where the sale of such numbers would break up a complete volume. Only sixty complete sets of the PROCEEDINGS are available for sale—Volumes I to XXX price $200.00. Orders should be sent to the Home Secretary, National Academy of Sciences, 2101 Constitution Ave., Washington 25, D. C.

The following publications are issued by the NATIONAL RESEARCH COUNCIL. Orders and inquiries should be addressed: Publication Office, National Research Council, 2101 Constitution Ave., Washington 25, D. C.

(A) THE BULLETIN OF THE NATIONAL RESEARCH COUNCIL, published at irregular intervals. Individual numbers priced variously. A discount of 20% from the list price will be given to all who subscribe for the series as issued, direct from the Publication Office.

(B) THE REPRINT AND CIRCULAR SERIES OF THE NATIONAL RESEARCH COUNCIL, individual numbers variously priced. A discount of 20% from the list price will be given to all who subscribe for the series as issued direct from the Publication Office.

PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 37 January 15, 1951 Number 1

ON THE ORIGIN OF THE SOLAR SYSTEM

By Gerard P. Kuiper

Yerkes Observatory, University of Chicago

Read before the Academy, October 12, 1950

Statement of the Problem.—A satisfactory theory of the origin of the solar system must account for the presence and the properties of the planets and the smaller bodies surrounding the sun, and preferably, but not necessarily, for the dynamical properties of the sun also. This means that we shall be concerned with the following bodies and properties:

(*1*) *Nine Planets:* The orbits around the sun are nearly coplanar and nearly circular; the largest inclinations and eccentricities are those of the outermost and innermost planets, Pluto and Mercury. The motions around the sun are all in the same sense (direct, by definition). The rotation of the planets is also direct, with obliquities less than 30°, except for Uranus where the angle has the exceptionally high value of 97°. The distances of the planets from the sun exhibit some degree of regularity (Bode's law). The masses of the four inner planets are roughly $10^{-6}\odot$ and the densities 4.1–5.5 cgs.; while the four Jovian planets have masses some hundred times larger and densities between 0.7 and 2.5 cgs.

(*2*) *Thirty Known Satellites:* The satellite systems vary from the beautifully regular case of Uranus to a completely irregular system like Neptune and the abnormal Earth-Moon system. Partly regular and partly irregular systems are those of Jupiter and Saturn. "Regularity" is measured by low inclinations with respect to the planetary equator; small orbital eccentricities, e; direct motion with respect to the planetary rotation; and some degree of regularity in the mean distances, a, to the planet.

(*3*) *The Asteroids and Meteorites:* The a-values of the 1168 largest asteroids[1] are roughly distributed according to $a = 2.89 \pm 0.24$ (*p. e.*) astr. units, though irregularities and fine structure occur in the distribution (Kirkwood Gaps) caused by the perturbations of Jupiter. Some 30,000 asteroids are accessible to telescopic observation and the indications are that their a distribution is not very different from that of the largest mem-

bers. Meteorites strike the Earth after having arrived in orbits showing a distinct similarity to those of asteroids ($i \simeq 10°$, $e \simeq 0.2$).

(4) *The Comets, Both Periodic and Non-periodic, with Their Associated Meteor Showers:* The i-values are nearly random except for the short-period comets and meteors which show a preference for small i and direct motion. The e-values are moderate for the short-period comets and almost unity for the remainder. Since the long-period comets with $e < 0.9$ cannot approach the Earth, we have no direct knowledge of them.

The sun has an angular momentum of only about 0.3% of that possessed by the entire planetary system. The smallness of this ratio requires an explanation, but the process controlling the origin of the planetary system is not necessarily the sole cause of it.

Analogy with Binary Stars.—The solar system is not the only multiple system known to astronomy. Simple calculation shows that planetary systems other than our own would be inaccessible to observation. But companions massive or luminous enough to be observable are found in great numbers. A detailed study of the nearest stars shows that at least half of these are in fact double or multiple systems.

The *mean separations* a in binaries are roughly given by[2] an error curve in $\log a$ having a dispersion in $^{10}\log a$ of $\sqrt{2}$:

$$F(\log a) = Ce^{-[(\log a - 1.27)/2]^2} \qquad (1)$$

The median value of a is therefore about 18 astr. units, a value remarkably close to the position of the most massive planets in the solar system.

The *mass ratios* in binary systems of approximately the solar mass have a distribution roughly given by[3]

$$f\left(\frac{m_2}{m_1}\right) = \text{constant} \left(\text{valid for } 0.3 < \frac{m_2}{m_1} < 1.0\right) \qquad (2)$$

The data for $0 < \frac{m_2}{m_1} < 0.3$ are still too sporadic to give a trustworthy check on equation (2) in that range. But several mass ratios down to about 0.02 have been found and equation (2) appears to give the correct order of magnitude of the relative frequency of such cases. The absolute frequency of binaries with large mass ratios is found by putting the constant in equation (2) at about 0.7, i.e., a quantity of the order of 1. We shall, therefore, obtain a tentative estimate of the occurrence of small companions by using

$$f\left(\frac{m_2}{m_1}\right) = 1. \qquad (2')$$

Anticipating the conclusion made below that the formation of the planetary system was a special case of the very general process of binary formation, we may estimate the probability of this event by means of equation (2′). As the appropriate mass m_2 we may use either the mass of Jupiter, $10^{-3}\odot$, or the hypothetical mass of the entire solar nebula from which all the planets formed, $10^{-1}\odot$. The use of the second quantity requires a further estimate of the fraction of cases in which a secondary nebula of mass $10^{-1}\odot$ was spread out and formed several small bodies instead of a single stellar companion. This fraction, obviously < 1, may as to order of magnitude be put at 10^{-1}. The probability that a main-sequence star like the sun is attended by a planetary system may therefore be estimated to be of the order of 10^{-2} or 10^{-3}. On the basis of these developments the event would be enormously more probable than on the basis of the Tidal Theory or any other theory requiring a stellar collision or near collision; the latter would make the probability per star of the order of $10^{-10^{1/2}}$, or 10^8 times smaller than according to the process envisaged here.

The problem of the origin of binaries is considered in a separate article;[4] in it a provisional interpretation is given of equations (1) and (2). It is shown that virtually all stars are likely to have originated as components of double and multiple systems. This follows from the result that the probability of a proto-star having a small-enough angular momentum to allow the formation of a single, rotationally stable, star is less than 10^{-3}. Many of the original double and multiple stars must have dissolved later due to internal gravitational perturbations and effects of passing stars, so that at present the fraction of apparently single stars is substantial, roughly $^1/_4$ to $^1/_3$.

The analogy between binaries and solar system, suggested by the similarity of internal dimensions, will be complete if it can be shown that a small pre-stellar cloud moving about the primitive sun may form either a single stellar companion or a family of planets, depending on the mass distribution within this secondary cloud.

The Solar Nebula.—The composition of the planets, particularly of the terrestrial planets, differs appreciably from that of the sun. It is now realized that the composition of the sun is typical for that of stars and even interstellar matter; it may therefore be called the *cosmic* composition. The cloud from which the planets formed will almost certainly have been of cosmic composition. A comparison of the planetary composition with the cosmic composition will therefore give a lower limit to the mass of the nebula from which the planets condensed and will contain many clues to the condensation process itself. In this manner the mass of the solar nebula may be estimated to be in excess of $0.01\odot$ and probably more nearly $0.1\odot$. The realization that the pre-planetary cloud (or *solar nebula*) had this comparatively large mass has caused a revision in our concept of the possible processes by which planetary condensation may have taken place.

It appears now that gravitational action of the nebula on itself can have been important. It is seen below that precisely this property has made the formation of planets possible at all.

We shall now sketch the principal processes leading up to the formation of the planets, their masses, orbits, rotations and satellites. A more complete account is given in two papers now in press.[5]

Celestial mechanics indicates that the planets have not altered their distances from the sun by large factors.[6] (Later we shall find that the alterations were, in fact, remarkably small.) The solar nebula must therefore have extended at least from Mercury to Pluto. It appears that the density distribution in this nebula need not be assumed, but can actually be derived from the present planetary distances; but as a first step we may simply assume that the total mass of the nebula is $0.1\odot$, as is suggested by the present planetary masses and compositions. This assumption had already been made by v. Weizsäcker[7] who found that turbulent dissipation will cause the solar nebula to dissolve with a half-life of about 10^7 years. The dissipation is caused by the different angular velocities of the different parts of the nebula. The time available for the formation of condensation products is therefore limited.

Another process, working in the opposite sense, is the slow collapse of the nebula parallel to the resultant axis of rotation (the z-coordinate). The initial stages of the collapse are fast, everywhere comparable to the orbital period of rotation around the sun. The later stages are much slower, 10^4–10^5 years in duration, and governed by the rate at which excess heat, caused by contraction, can be radiated away. Therefore, a *maximum of density* will result, some 10^4–10^5 years after the solar nebula got first in place; it is expected that the planets segregated close to or somewhat before that epoch.

Condensation products will form wherever the vapor is supersaturated.[8] Unless some large-scale coalescing property within the solar nebula can be found we must anticipate the formation of enormous numbers of small condensations (a process somewhat comparable with the formation of raindrops in a cloud), each growing slowly with time. The coalescing property needed to account for a small number of large planets was assumed by Weizsäcker to be a regular turbulence pattern in the solar nebula. Weizsäcker assumed a geometrical pattern of primary vortices, separated by rings of secondary "roller-bearing" vortices; he further assumed that condensation would progress most rapidly along the roller-bearing circles. This was assumed to lead to masses of sufficient prominence to insure their stability against subsequent destruction and hence their subsequent growth by accretion of neighboring material. Ter Haar[9, 10] elaborated Weizsäcker's picture and refined the discussion of the condensation process. Chandrasekhar and Ter Haar[10] showed that the regular turbulence pattern

envisaged by Weizsäcker would not arise but should be replaced by the Kolmogoroff spectrum of turbulence previously derived for non-rotating media. The writer thereupon showed[5] that secondary eddies were less suited to foster condensation than the primary eddies of the Kolmogoroff spectrum, because any "roller-bearing" eddies that may form when large clouds collide would be short-lived and of distinctly higher temperature, but only slightly higher density, than the primary eddies. He further showed that even the primary vortices had comparatively short lifetimes, of the order of 10^2 years rather than 10^6 years as is necessary for the formation of appreciable condensations. The repeated formation and dissolution of primary vortices in more or less random positions woud lead to innumerable small condensations throughout the solar nebula instead of a few large ones. This obstacle can be overcome only if it is assumed that gravitational action intervenes and keeps some of the large vortices together long enough for massive condensations to form. This requires that these vortices have a density large enough to protect them from tidal disruption by the sun, i.e., their density must exceed the critical value, ρ_R, at which the internal attraction within the cloud and the tidal force by the sun are equal. This critical density is called the *Roche* density; it is intimately related to the well-known Roche limit, used, e.g., in the discussion of the stability of satellites and the Ring of Saturn. The Roche density varies inversely as the cube of the distance from the sun and is given by

$$\rho_R = \left(\frac{2\beta R_\odot}{a}\right)^3 \rho_\odot = \frac{6\beta^3 M_\odot}{\pi a^3} = \frac{24\pi\beta^3}{GP^2}, \qquad (3)$$

in which β is a quantity of the order of unity, a the distance to the sun and P the corresponding period of revolution around the sun; $R_\odot$ and $M_\odot$ are the radius and mass of the sun. The value of ρ_R is about 10^{-5} cgs. near Mercury, 10^{-6} near the Earth, 10^{-8} near Jupiter and 10^{-10} near Neptune. (The precise values depend not only on the mass and the distance to the sun but also on the molecular and turbulence velocities;[5] the range of uncertainty caused by the latter is covered by the approximate nature of β.) Condensation of masses of planetary size could not have occurred unless the mean density of at least some vortices or clouds in the solar nebula exceeded ρ_R. On the other hand, the mean density of these clouds could not have exceeded ρ_R by a large factor either. Such an excess in density would have caused the domain of the cloud to extend beyond its initial boundary and therefore the adjacent parts of the solar nebula would automatically have become part of the cloud until the mean density was lowered to a value of the order of ρ_R itself.

The conclusion to which we have come, that the segregation of pre-planetary masses took place essentially at the Roche density, appears to be capable of empirical verification. It can be shown[5] that the stability of

two small spherical masses in contact, each at the Róche density, is not appreciably affected by the presence of the remainder of the solar nebula except near its outer edge. Therefore if m and r are the masses and the radii of the small spheres, we find from (3):

$$\frac{m}{M_\odot} = \left(\frac{2r\beta}{d}\right)^3. \tag{4}$$

If we introduce $\Delta = 2r = a_2 - a_1$, the distance between the centers of the masses whose mean distance from the sun is a, we have

$$\frac{m}{M_\odot} = \left(\frac{\beta\Delta}{a}\right)^3. \tag{4'}$$

If Δ should in some cases be as large as $a/3$ or $a/2$, as is indicated by the present planetary separations, we must allow for the fact that the flat disk-shaped solar nebula would not have filled such large spheres except for an equatorial slice occupying the fraction f of the order of perhaps 10^{-1}. Further, a comparison of the cosmic abundances with the present chemical composition of the planets shows that no more than a fraction g of the original pre-planetary clouds has been retained, g being about 10^{-1} for Jupiter and 10^{-3} for the Earth. The geometric mean of g for all planets may be roughly 10^{-2}. Therefore, if M_P represents the present planetary masses, we should expect

$$\frac{M_P}{M_\odot} = fg\left(\frac{\beta\Delta}{a}\right)^3, \tag{4''}$$

while empirically we find approximately[5, 11]

$$\frac{M_P}{M_\odot} = 10^{-4}\left(\frac{\Delta}{a}\right)^3. \tag{5}$$

It may be shown[5] that f is between 10^{-1} and 10^{-2}. The remarkable agreement between the predicted relation (4'') and the empirical relation (5) *shows that the segregation of planetary masses did take place essentially at the Roche density.* The closeness of the agreement is appreciated by remembering that the Roche density varies by the factor 500,000 between Mercury and Neptune and the M_P values are proportional to this quantity.

It appears that the comparison between equations (4'') and (5) may even be refined to take account of the variation with M_P of g in (4''). For the pair Jupiter-Saturn the value of f is about $10^{-11/3}$, while g is roughly 10^{-1}. If $\beta = 1$ we should for this pair expect the coefficient in (5) to be $10^{-21/3}$. This is precisely what is found.[11] For the pair Earth-Venus, $f = 10^{-1}$, $g = 10^{-3}$, roughly, and hence the expected coefficient in (5), 10^{-4}. This, again, is confirmed.[11] It is significant that the empirical relationship (5) holds for both planets and satellites; we return to this comparison later.

These examples show that the systematic deviation[11] from equation (5) is caused by the systematic change of g with M_P. A second result,[11] that pairs with very unequal masses are systematically somewhat displaced with respect to pairs of nearly equal masses, has not yet been interpreted.

With the density in the plane of symmetry determined, the density distribution at right angles to this plane (z-direction) may be computed. Weizsäcker and Ter Haar in their discussions of the hydrodynamics of the solar nebula assumed that the z-component of the force of gravity is given by the z-component of the solar attraction. With the equatorial density as high as ρ_R the z-force near the plane of symmetry will in fact be predominantly caused by the attraction of the nebula itself, in much the same way as the z-force in the galaxy near the galactic plane is determined by the

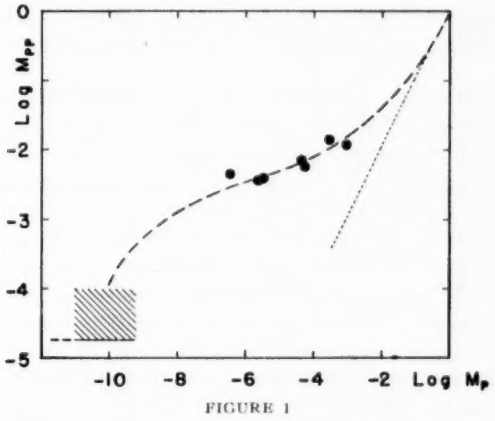

FIGURE 1

Relation of computed mass of proto-planet, M_{PP}, to present observed mass of planet, M_P. Unit of mass is Sun.

density distribution of the surrounding parts of the galaxy and not by the galactic nucleus. The ratio of the two z-forces, nebula/sun, is $24\beta^3$, or about 25. The solar nebula is flatter than considered by previous investigators,[7, 9, 10] by this same factor. The resulting large density increase in the new model allows the operation of *gravitational instability*. It is this mechanism which amplifies initial density fluctuations caused by turbulence[5] and is responsible for the formation of large clouds (proto-planets) which are gravitationally stable. These stable clouds exist long enough for the condensation processes to run their course and for large masses to form instead of billions of small, unrelated drops or flakes.

The density distribution in the z-coordinate may be computed from an equation resembling the barometric formula, the field of force being known

and also, approximately, the molecular velocities. The latter follow from the distribution in temperature with distance from the sun, which in turn may be estimated when allowance is made for the opacity of the nebula.

The condition that the proto-planets segregated at the Roche density, together with the density gradients in z just found, fixes completely the mass distribution for each proto-planet, both in the xy plane and the z-direction. We need only to add the dimensions in the xy plane for a complete description of the proto-planets, including their masses. The diameters of the proto-planets, at their maximum extent, are found by assuming that the proto-planets grew until they touched their neighbors on each side and by the further assumption that the planets formed at their present

TABLE 1

PLANET	LOG a	LOG a BOUNDARY	Δ LOG a	a BOUNDARY	MAX. $2R(PP)$	MAX. $M(PP)$	LOG $\frac{M(PP)}{M(P)}$
Mercury	−0.330		....		....	....	....
(aph.)		−0.24		0.58			
Venus	−0.141		0.17		0.27	0.0037	3.18
		−0.07		0.85			
Earth	0.000		0.16		0.38	0.0038	3.10
		+0.09		1.23			
Mars	+0.183		0.18:		0.63:	0.0045:	4.14:
		+0.27:		1.86:			
Asteroids	+0.45:		0.28:		....	....	....
		+0.55:		3.55:			
Jupiter	+0.716		0.30:		3.5	0.012	1.0
		+0.85		7.1			
Saturn	+0.979		0.30		7	0.014	1.69
		+1.15		14			
Uranus	+1.283		0.23		10	0.0072	2.22
		+1.38		24			
Neptune	+1.478		0.20:		14	0.0057:	2.04
		+1.58:		38			
Pluto	+1.597		....		....	....	....

NOTE: Unit of a is a. u. PP = proto-planet; P = planet. Cf. figure 1.

distances to the sun. The resulting partitioning of space is shown in table 1; it was done in a logarithmic plot because of the approximate validity of Bode's law and the fact that all relevant functions have an exponential run with a. The hypothesis that the proto-planets at their maximum size touched each other, accounts for the fact that at that point further growth was arrested; nearly all of the solar nebula would thus have been "swept clean." At the same time a mechanism is provided for the explanation of retrograde satellites in the outer regions of the "sphere of action" of such planets as did not lose thereafter too great a fraction of their original mass (Jupiter and Saturn).

The preceding arguments fix the masses of the proto-planets, M_{PP}, uniquely. They are found in table 1 and plotted against the present planetary masses, M_P, in figure 1. The closeness of the relation is very remarkable. The largest deviation is that of Mars, amounting to $10^{0.2}$ or a factor 1.6 in M_{PP}. This may be interpreted to mean that the diameter of proto-Mars was in fact only 0.8 of that assumed; this difference is not serious since the outer boundary of the Mars zone is ill-defined.

The curve relating M_{PP} to M_P has been tentatively extended beyond the region of the planets, as follows. For *any* value of M_P we must have $M_P \leq M_{PP}$. The limiting line so defined is dotted in figure 1; very probably this line is an asymptote. Further, it has been shown[5] that there is a theoretical lower limit to M_{PP}, given by $\log M_{PP} = -4.8$. All points must necessarily lie above this limit. It is assumed that Ceres and a few other asteroids are true planetary condensations. Their M_P values are about $10^{-10}\odot$. The shaded region in figure 1 shows the most probable location of these objects.

The closeness of the relation in figure 1 shows two things: (*a*) The position of Uranus and Neptune is intermediate between that of the Earth and Venus on the one hand and that of Jupiter and Saturn on the other. This shows that the relation in figure 1 is not strongly affected by the distance to the sun, though an extra factor of two or three in the present masses of the Jovian planets would hardly be discovered in this manner. This means, e.g., that the Earth is different from Jupiter *not* because it originated closer to the sun but primarily because proto-Earth was less massive than proto-Jupiter. The existence of a planet like Pluto outside Neptune becomes at once "acceptable." (*b*) The distances of the planets to the sun cannot have greatly changed during the last 3.10^9 years, at least not in an irregular fashion. If they had, the relation in figure 1 could not possibly have been close since M_{PP} depends in a sensitive way on the present planetary distances. This conclusion goes much beyond the powers of celestial mechanics which for such extremely long intervals has made only the most general predictions.[6]

The masses of the proto-planets are about as large (within 20 or 30%) as might have been expected from the present planetary compositions. This conclusion is an important check on the theory presented. It is the first time that planetary masses have been predicted with any degree of accuracy.

By allowing $0.003\ M_\odot$ for each of the proto-planets which formed Mercury, the asteroids and Pluto we find from table 1 the total mass of that part of the solar nebula that was responsible for the formation of the planets. It is found to be

$$\sum M_{PP} = 0.060\ M_\odot. \qquad (6)$$

The total mass of the solar nebula must have been somewhat (but probably not appreciably) larger and may be put in round figures at $10^{-1}\ M_{\odot}$.

The Bode Law. Planets and Binaries.—We must now consider the factors that determine the relative spacings, Δ/a, of the planets. Why did nine planets form and not 10^2 smaller ones?

The mode of formation of the proto-planets gives a clue. If the model of one proto-planet per ring, valid for the large planets, is extended to smaller and smaller bodies, we note that a smaller and smaller fraction of the ring is occupied to the Roche density. It follows that the spacing of the planets at the distance a from the sun is determined by the ratio of the density in the solar nebula to the local Roche density.

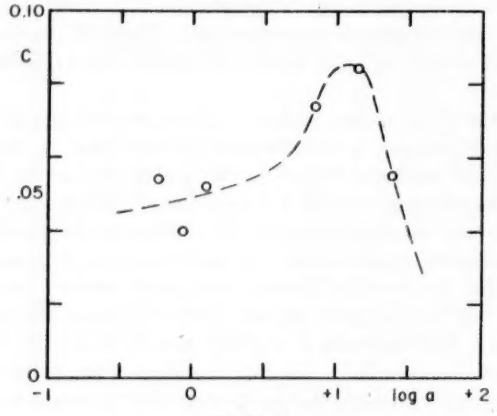

FIGURE 2

Mass distribution in solar nebula prior to formation of proto-planets (tentative).

In this manner the density distribution is found which existed in the solar nebula prior to the break-up into proto-planets. In order to express the result we take as reference a nebula having everywhere the Roche density in the xy plane. Such a nebula will have a given "surface density" of matter when projected on the xy plane. In fact, it is one in which the amount of matter per unit interval of log a is constant. In terms of this standard surface distribution, the pre-planetary nebula had the density 0.05 from about $^1/_2$ to 2 astr. units, rising to 0.09 at 10 astr. units, and falling off to below 0.05 at 30 astr. units. It is illustrated in figure 2.

From the theory of gravitational instability it follows that a lower limit exists below which stable proto-planets cannot form. This limit in the units of figure 2 is found to be

$$C_0 = 0.05\alpha^{-1}\beta^{-3/2}, \qquad (7)$$

in which both α and β are quantities of the order of unity. It appears therefore that the solar nebula had only just enough surface density to cause formation of planets. A reduction by a factor of two or three (or more) would have been fatal to this process; billions of comet-like bodies would have formed instead. This result may also be interpreted to mean that the hydrogen content of the solar nebula could not have been much less than supposed; a reduction by a factor three over the adopted cosmic composition would have caused a nearly equal reduction in mass, and would again have prevented planets to form. It is possible that the hydrogen content was actually somewhat larger than assumed.

If, on the other hand, the surface density of matter over the xy plane had been larger than found in figure 2 by a factor of three or more, no ordinary planets would have formed either but instead one or more stellar companions. This results from the circumstance that a rather small increase in C (see Fig. 2) causes a large increase in the mass of the proto-planet, which in turn causes a still greater increase in the mass of the final body (Fig. 1).

Three conclusions follow: (*a*) the so-called Bode law is no real law but a consequence of the density distribution in a which the solar nebula happened to have; the larger the density, in terms of the local Roche density, the larger the ratio between the radii of two consecutive orbits and the larger the masses of the proto-planets (and, finally, the larger the masses of the planets); (*b*) the process of formation of a planetary system is a special case of the almost universal process of binary-star formation; the crucial parameter is the surface density of the secondary nebula in terms of the local Roche density; (*c*) inspection of figure 1 suggests that the subsequent evolution of the proto-planets is not greatly influenced by the distance to the sun, but instead largely determined by the mass of the proto-planet itself (once formed, the object is "on its own"). Space does not permit to discuss the theoretical interpretation of this conclusion; it appears to be connected with the rate at which the lighter gases can escape.[5]

Conclusions.—We may now turn to the problems listed on page 1 and list the solutions now at hand or indicated. The common direction of revolution and the low relative orbital inclinations are accounted for by the flatness of the solar nebula. The internal viscosity of the nebula accounts for the near-circular orbits. The fact that both Mercury and Pluto are exceptions in their inclinations and eccentricities may be attributed to the absence of constraining action on the proto-planets formed on the fringes of the solar nebula. The direct rotation of the planets is attributed[5] to solar tidal friction on the proto-planets. The solar tidal force nearly equals the self-attraction for each of the proto-planets at their maximum extension; i.e., in the proper units Neptune is no farther away from the sun

than Mercury and the tidal effects are equally large in both cases. Regardless of any initial rotational motion a direct rotation will be forced upon the proto-planet, with a period equal to the orbital period. As is readily verified, this leads to an amount of angular momentum per unit mass some 10^4 times greater than found on the present planets. Part of this is lost during the evaporation process of the proto-planets (the ejected molecules carry off more than the average amount of angular momentum per unit mass); while part of it is lost by continued solar tidal friction during the contraction process. The latter cause has a secular effect on the obliquities; it has been shown[5] that they will increase some three or fivefold, from initial obliquities of the order of 3° (expected from the turbulent solar nebula, and consistent with the relative orbital inclinations) to the present values. The largest obliquity to which this process can lead is 90°; retrograde rotation cannot arise by the processes considered. It is not clear why Uranus has passed the upper limit by 7°; possibly some extraneous object has moved through the solar system. The present periods of rotation have not yet been accounted for quantitatively. This appears to be a very complex problem, with physics, chemistry and dynamics all playing a role. We have here perhaps the most important potential source of information still unused in the reconstruction of the planetary condensation processes.

The regular satellites may be explained in a manner analogous to that found for the planets themselves.[5] Little progress has been made so far with their condensation processes, which should prove very instructive in view of the large density differences known to exist among the satellites. The retrograde satellites of Jupiter and Saturn have been interpreted[5] as having been caused by glancing collisions between the corresponding proto-planets. They were assumed to have been retained by these large planets only because these planets lost a much smaller fraction of their initial mass. It is possible, however, that capture has played a role instead. This requires further investigation. The asteroids were not formed in a region of low density in the solar nebula. In such a region no planets of any kind could have formed. Rather we must assume that the density was well above C_0 of equation (7), but that the formation of a normal-size proto-planet was prevented by proto-Jupiter (mass = $0.012\odot$). It can be shown that in the presence of strong perturbations a small proto-planet, of a given density close to the local Roche density, is more stable than a large one of the same density. The total number of small proto-planets estimated to have formed in the region between Mars and Jupiter is between 5 and 10. They formed small planets, like Ceres (cf. figure 1 and accompanying discussion). It is assumed that two of these collided sometime during the last 3.10^9 years, an event having a sufficiently large probability. Thereafter secondary collisions became increasingly frequent. The most

recent of these collisions account for the Hirayama families. In this manner thousands of asteroids were formed, being the largest of the fragments, as well as billions of meteorites.[12]

The outermost region of the solar nebula, from 38 to 50 astr. units (i.e., just outside proto-Neptune), must have had a surface density below the limit set by equation (7). The temperature must have been about 5–10°K. when the solar nebula was still in existence (before the proto-planets were full grown), and about 40°K. thereafter. Condensation products (ices of H_2O, NH_3, CH_4, etc.) must have formed, and the flakes must have slowly collected and formed larger aggregates, estimated to range up to 1 km. or more in size. The total condensable mass is about 10^{29} g., but not all of this could be collected. These condensations appear to account for the comets, in size,[13] number[13] and composition.[14]

The planet Pluto, which sweeps through the whole zone from 30 to 50 astr. units, is held responsible for having started the scattering of the comets throughout the solar system. Pluto's perturbations will have caused initial, near-circular, cometary orbits to become moderately elliptical; thereupon stronger perturbations by Neptune and the other major planets will have scattered them even more broadly. As Oort[13] and others have shown, the quantity which is spread nearly uniformly in both directions is the quantity a^{-1}, the reciprocal of the semimajor axis (which is related to the energy of the object). A certain fraction of the comets will be scattered in the region of very small a^{-1} values, i.e., in the outer regions of the "sphere of action" of the sun. As Oort[13] has shown, stellar perturbations will redistribute the orbital elements there, and in particular make the motion around the sun one of random orientation. Oort[13] shows that the dynamical half-life of a comet in this outer region is about 10^{10} years. The comets which we observe today were sent back to the inner regions of the solar system by small random stellar perturbations. The above views are an adaptation of Oort's[13] dynamical analysis; but we differ in our hypothesis as to the region where the comets originated. Oort[13] assumes that they were formed between Mars and Jupiter, in association with the origin of asteroids. The composition of the comets indicates condensation at a very much lower temperature, around 10°K., consistent with the region of origin proposed here. The evaporation and subsequent complete disintegration of comets into the minute particles which cause meteors and the Zodiacal Light is also understandable from their formation outside Neptune. Asteroidal bodies would be expected to remain intact or possibly break up into a few large fragments.

The theory described here does not depend on any specific *ad hoc* assumptions. Certain assumptions which were made at the outset, e.g., that the planetary distances have not changed appreciably or that the solar nebula was approximately of cosmic composition, appeared capable of verification

afterwards. One assumption, that the sun was already formed as a star and of a luminosity approximately equal to that found today, requires further study.[15] Certain investigations on the contraction and condensation process of the proto-planets need still be made, including the analysis of solar tidal friction on these composite structures. Finally, the cause of the small solar rotation must be cleared up; it is undoubtedly connected with the larger problem of why nearly all *G*-type dwarf stars, in single and in binary systems, have such slow rotations. It is felt, therefore, that this problem is not necessarily a part of a theory on the origin of the solar system.

The probability of a star being attended by a planetary system was estimated to be between 10^{-2} and 10^{-3}. The total mass of the galaxy is about $2.10^{11}\odot$; while the average stellar mass is about $0.5\odot$. From these figures the total number of planetary systems in the galaxy is estimated to be of the order of 10^9. One can only speculate on the possible forms of life which may have developed on these many unknown worlds.

[1] Those with $g \leq 10.9$ in the 1947 Catalogue of Minor Planets. The quantity g is the magnitude of the asteroid reduced to unit distance from the Earth and unit distance from the Sun; it depends chiefly on the *size* of the body since the albedos are probably not very different.

[2] Kuiper, G. P., *Publ. Astr. Soc. Pac.*, **47,** 138 (1935).

[3] *Ibid.*, p. 147.

[4] Kuiper, G. P., *Astrophys. J.*, in press, (1951).

[5] Kuiper, G. P., "On the Origin of the Solar System," Chap. 8 of *Astrophysics*, edited by J. A. Hynek, McGraw-Hill Book Co. (in press). (This paper was submitted for publication November, 1949; it was given a limited distribution in February, 1950.) Also, *Ibid., Astrophys. J.*, 1951.

[6] Brown, E. W., *Publ. A.S.P.*, **44,** 37 (1932).

[7] Weizsäcker, C. F. v., *Zeitschr. f. Astrophys.*, **22,** 319 (1944).

[8] Becker, R., and Doering, W., *Ann. Phys.*, **24,** 719 (1935). Ter Haar, D., *Bull. Astr. Inst. Neth.*, **10,** 3 (1943).

[9] Ter Haar, D., "Studies on the Origin of the Solar System," *Kgl. Danske Videnskab. Selskab, Mat-Fys. Med.*, **25,** No. 3 (1948). A summary appeared in Science, **107,** 405 (1948)

[10] Ter Haar, D., *Astrophys. J.*, **111,** 179 (1950). Chandrasekhar, S., and Ter Haar, D., *Ibid.*, p. 187.

[11] Kuiper, G. P., "The Law of Planetary and Satellite Distances," *Ibid.*, **109,** 308 (1949).

[12] Kuiper, G. P., "On the Origin of Asteroids," *Astron. J.*, **55,** 164 (1950).

[13] Oort, J. H., *Bull. Abstr. Inst. Neth.*, **11,** 91 (1950).

[14] Whipple, F. L., *Astrophys. J.*, **111,** 375 (1950).

[15] Actually, it is more probable that the Sun had not completed its contraction and that its radius was very substantially larger than at present (by a factor in excess of 10 and possibly as large as 50). This will have caused a reduction in the solar luminosity by the square root of this ratio, say 5. This reduction in solar temperature facilitates the processes considered in the text.

PROXIMA CENTAURI AS A FLARE STAR

BY HARLOW SHAPLEY

HARVARD COLLEGE OBSERVATORY

Communicated November 11, 1950

1. The dwarf red companion of the bright double star Alpha Centauri was detected by Innes in the course of an examination of plates made with the Franklin Adams camera at the Union Observatory, Johannesburg, in 1915.[1] Although the object was two degrees distant from Alpha Centauri, its similar motion, and subsequently its comparable parallax, clearly showed the physical association of the three objects. The somewhat larger proper motion and the slightly larger parallax, derived by some of the observers, have led to the conclusion that Proxima Centauri is actually a little nearer than Alpha Centauri, and is therefore the nearest of all stars so far discovered; hence the appropriateness of its name. The difference in the parallax, however, scarcely exceeds the probable error of the determinations, and van de Kamp gives in his listing of the stars nearer than five parsecs[2] identical values for the parallax, 0".761, and for the annual proper motion of Alpha, 3".68, and of Proxima, 3".85.

2. It became clear some years ago that such low-luminosity red dwarfs, at the faintest end of the main sequence, are not simple and quietly perishing stars. Castor C, for example, was found at Mount Wilson to be a close pair of dwarf Class M stars. It is an eclipsing variable with a period of 0.814 days. Later, in the course of his parallax and proper motion work, van Maanen found that Lalande 21258B and Ross 882 had each shown one abnormal brightening among the many images appearing on the Mount Wilson plates.[3]

In 1949, Luyten's remarkable double, L 726-8, discovered by him as a large proper motion star on the Harvard Bruce plates, was found to be a light variable. Not only is it one of the nearest of all stars, and among the dimmest stellar radiators yet on record, but also the fainter component of the pair showed brief flares on two occasions. The star has been fully discussed by Luyten.[4] In their concurrent work on the spectra, Joy and Humason[5] also found a flare, especially strong in blue light. Subsequently 569 Harvard plates yielded seven additional flares,[6] one of them corresponding to a momentary increase of forty times in the light intensity. The results for L 726-8 not only gave support to van Maanen's earlier observations, but attracted general attention to the dwarf red stars. There are not many of them on record, and they cannot be seen very deep in space; but here were three affected by temporary flashes—a novelty in the photometry of stars. The spectrum of Luyten's star showed unusual charac-

teristics, including bright calcium and hydrogen lines.[5] Obviously these red stars are not cooling off monotonously.

The list was increased to four by the interesting photoelectric discovery and observation by Gordon and Kron of an hour-long flaring-up of the M dwarf, BD + 20°2465, which confirmed earlier evidence for the other flare stars that the duration of the spurt and recovery in brightness is indeed of short duration.[7]

3. Before any of the foregoing information on flare stars had become known, nearly a hundred Harvard photographs of the field of Proxima Centauri had been examined in search of variability of the continuous sort that affects the dwarf red variable of the Castor system. The results were negative. The constant photographic magnitude was established, with the aid of the most convenient magnitude sequence, as 13.4, with small dispersion around that value except for one very discordant measure which was ignored. When Thackeray recently published a discussion of the variable spectrum of Proxima[8] and reported also that the observers at the Cape Observatory had examined 58 plates of the star, taken for parallax during the past two decades, and found no certain evidence of flares, the earlier Harvard measures were reexamined and the one most discordant observation was recognized as a probable flare. Four measures of 13.2, 13.2, 13.0, and 13.2 possibly also indicate flare phenomena, and are included in the later tabulation. The plates had been made with the 10-inch Metcalf on a neighboring Milky Way Field, No. 171, and Proxima is in a rather unsatisfactory position on the photographs.

An examination of several series in the Harvard plate collection has now been undertaken. The MF plates were searched by Miss C. D. Boyd and the other series by Mrs. V. McK. Nail. Without going further back than 1925, five hundred additional plates were found on which the image of Proxima Centauri appears. The following tabulation gives the plate series, name of telescope, the average exposure time, the number of plates examined, and the number of times that Proxima Centauri was found to be of magnitude 12.8 or brighter.

MF	Metcalf triplet	45 min.	93 plates	1 flare
B	Bache doublet	45	80	6
RB	Ross-Fecker lens	90	170	15
AM	Cooke lens	90	211	19
AX	Tessar lens	120	35	7
A	Bruce doublet	120	3	0
	Total		592	48

It appears that on eight per cent of the plates (six per cent of the exposure time) the star was brighter than average by a distinctly measurable amount, making it the most active flare star on record. By going to the older

Harvard plates the list of recorded flares could be substantially increased. Possibly such a study should be made if it later appears that there is an indication of periodicity in the outbursts, but at present it is more important to follow some of these flares carefully with photoelectric photometers or short exposure photographs in order to learn more about the nature of the outbursts than can be obtained from these long-exposure Harvard plates. Proxima has now been put on the program of the photoelectric photometer at the Harvard station in South Africa.

For all series the exposures are so long that brief sharp maxima are much smoothed out. The brightest recorded maxima are 12.4, indicating amplitudes of at least one magnitude. The AX camera has a steep color curve and is not very reliable for work on such red stars.

On three nights consecutive exposures were made throughout 5.5, 2.25, and 7 hours with the 10-inch Metcalf telescope. No measurable change of brightness was shown. In addition, on seven nights plates with different instruments are available. Four pairs occur at times of flares; they give magnitudes consistent with the presumption of outbursts of an hour or so.

The Julian Day and decimal of a day for the flares on MF, B, RB, AM, and AX plates, with magnitude estimates by Mrs. Nail, are as follows:

JD AND GMT	SERIES	MAG.	JD AND GMT	SERIES	MAG.	JD AND GMT	SERIES	MAG
2423996.534	AX	12.6	2428040.215	B	12.6	2431184.427	AM	12.7
24244.827	AX	12.5	28276.553	RB	12.8	31197.562	B	12.8
24286.694	AX	12.4	28283.542	B	12.8	31311.343	AM	12.6
24294.693	AX	12.7	28289.417	AM	12.4	31330.264	AM	12.8
25446.285	AX	12.4	28896.552	AM	12.3	31494.591	AM	12.5
26091.372	MF	12.6	28994.495	AM	12.6	31584.310	B	12.7
26091.422	RB	12.9	29051.408	AM	12.8	31614.414	AM	12.8
26131.331	MF	13.2	29131.238	RB	12.8	31620.287	AM	12.8
26181.242	MF	13.2	29834.345	AM	12.4:	31873.323	AM	12.4
26189.286	RB	12.4	29850.233	RB	12.8	31887.529	AX	12.6
26215.220	MF	13.0	30134.463	AM	12.8	31921.424	AM	12.8
26812.531	RB	12.8	30474.446	RB	12.8:	31935.410	RB	12.8
27519.505	AX	12.4	30491.354	AM	12.6	32033.290	AM	12.7
27525.635	MF	13.2	30497.277	B	12.7	32324.277	B	12.6
27630.224	RB	12.6	30515.448	RB	12.6	32387.291	AM	12.6
27958.469	RB	12.8	30578.289	AM	12.7	32406.227	AM	12.6
28010.320	RB	12.8	30850.451	RB	12.8	33033.577	RB	12.8
			30900.320	RB	12.4			

4. A sixth star of spectral Class M, which has the spectrum characteristics of SS Cygni, may be assignable to this small group of flare stars. It is AE Aquarii, which has been studied at various observatories and is now on the program of the American Association of Variable Star Observers.[9] Recent unpublished observations show rapid changes in the course of an hour or so. Like Proxima Centauri, it merits study with photoelectric photometers and high speed spectrographs.

5. Dwarf red flare stars may become of considerable importance in considerations of stellar evolution. Something near one-half of the fifty nearest stars now known are dwarfs of Class M, with absolute photographic magnitudes fainter than +12. Of these, about twenty per cent are now known to be affected with sudden bursts of radiation, presumably emanating, as Gordon and Kron, Whipple, Greenstein[10] and others have pointed out, from relatively small areas which for a time explode with highly intense (blue) radiation, and quickly cool to the average conditions.

The great contrast in surface brightness between the intense blue outburst and the relatively cool surrounding surface greatly increases the total light of the red dwarf stars, whereas similar phenomena on the surface of the sun, or on stars of earlier spectral class, would not appreciably affect the total radiation. The total energy of an outburst on Proxima is equivalent to that of some of the solar flares currently observed; and although the solar flares do not measurably increase the magnitude of the sun, they do succeed in disturbing the earth's ionosphere.

6. Is it likely that measurable eruptive flaring disturbs the radiation of all the red dwarf stars? Possibly the phenomenon is universal, especially for red dwarfs with emission spectra, but the examination of 225 Harvard plates of the RB, RH and BM series on Barnard's Star (spectrum M 6, absolute magnitude +13.4, vis.) reveals no variability.

[1] Innes, R. T. A., *Union Obs. Circ.*, **30** (1915), and **40** (1917).
[2] van de Kamp, P., *Pop. Ast.*, **48**, 297 (1940).
[3] van Maanen, A., *Mt. Wilson Contr.*, **630** (1940); *Ap. J.*, **91**, 503 (1940); *Pub. Astr. Soc. Pac.*, **57**, 216 (1945).
[4] Luyten, W. J., *Ap. J.*, **109**, 532 (1949).
[5] Joy, A. H., and Humason, M. L., *Pub. Astr. Soc. Pac.*, **61**, 133 (1949).
[6] *Harv. Ann. Card.*, 1010 (July 1949).
[7] Gordon, K. C., and Kron, G. E., *Pub. Astr. Soc. Pac.*, **61**, 210 (1949).
[8] Thackeray, A. D., *Mon. Not.*, **110**, 45 (1950).
[9] Campbell, L., *Pop. Ast.*, **55**, 557 (1947).
[10] Greenstein, J., *Pub. Astr. Soc. Pac.*, **62**, 160 (1950).

THE THEORY OF MICRO-METEORITES.* PART II. IN HETEROTHERMAL ATMOSPHERES

BY FRED L. WHIPPLE

HARVARD COLLEGE OBSERVATORY

Communicated by Harlow Shapley, October 16, 1950

Introduction.—In a previous paper[1] the writer presented a theory for the passage of a small meteoritic body through the earth's atmosphere without damage by melting or vaporization. The theory, which derives the maximum ratio of mass to effective surface area for a micro-meteorite of a given initial velocity, is valid for an atmosphere of constant temperature and mean molecular weight but depends upon certain assumptions. The present paper extends the theory to more general atmospheric circumstances and also explores the tenability of the basic assumptions. For convenience, the previous notation is utilized while references and equations of the first paper are identified by asterisks.

1. The Solution in an Atmosphere of Constant Temperature Gradient.—At altitudes in the neighborhood of 100 kilometers, most investigators at present tend to adopt an atmospheric temperature increase with altitude.[2] Little is certainly known about the mean molecular weight at these altitudes but a constant value will probably be a fair assumption. The dissociation of O_2 probably occurs near 100 kilometers altitude but its effect on the density may be absorbed in the assumed temperature gradient for constant molecular weight, as may the slight effect arising from the decrease of gravity with height.

It is, therefore, of great interest to repeat the solution for the maximum ratio of mass to surface area, applying to micro-meteorites in an atmosphere with a constant temperature gradient. Let the atmospheric temperature, T_a, vary with height, h, according to the relation

$$T_a = T_1 + \nu h, \tag{1}$$

where T_1 is the temperature at a height, $h = 0$ and where ν remains constant.

The atmospheric density then can be expressed by the relation[3]

$$\rho = \rho_1 \left(1 + \frac{\nu h}{T_1}\right)^{-(1 + q/\nu)}, \tag{2}$$

where ρ_1 is the density at $h = 0$.

The quantity q is defined by

$$q = \mu g / k, \tag{3}$$

where μ is the mean molecular weight of the air, g is the attraction of earth gravity corrected for altitude and k is the Boltzmann Constant. The logarithmic density gradient, b, used previously is equal to $(q + \nu)/T_a$.

With our previously adopted assumptions that the drag coefficient, D, accommodation coefficient, α, and emissivity, β, are constants we may substitute ρ from equation (2) for its isothermal value in equation (14*) and integrate the velocity differential to obtain

$$\log (V/V_\infty) = -\frac{ADT_1\rho_1}{2mq\cos Z}\left(1 + \frac{\nu h}{T_1}\right)^{-q/\nu}. \qquad (4)$$

Equation (10*), giving the surface temperature of the micro-meteorite, now becomes

$$T_s^4 - T_0^4 = \frac{\alpha A \rho_1 V^3}{2\beta B\sigma}\left(1 + \frac{\nu h}{T_1}\right)^{-(1 + q/\nu)}, \qquad (5)$$

where we again neglect the heat capacity.

Upon eliminating the terms in h from equations (4) and (5), we express the temperature of the micro-meteorite in terms of the velocity, independently of height or density, as follows:

$$T_s^4 - T_0^4 = \frac{\alpha V^3}{\beta B\sigma}\left(\frac{2}{A\rho_1}\right)^{\nu/q}\left[-\frac{mq\cos Z}{DT_1}\log (V/V_\infty)\right]^{(\nu+q)/q}. \qquad (6)$$

By equating to zero the derivative with respect to V of the right member of equation (6) we find that the maximum temperature of the meteoroid, T_m occurs at the critical velocity, V_c, given by

$$\log (V_c/V_\infty) = -\frac{\nu + q}{3q}. \qquad (7)$$

This value of the velocity is now substituted into equation (6), A is equated to A_1 and expressed in terms of B by equation (2*) and the ratio of mass to surface area for the micro-meteorite is obtained in the form

$$\frac{m}{B} = \frac{3eDT_1}{(\nu + q)\cos Z}\left(\frac{\rho_1}{8}\right)^{\nu/(\nu+q)}\left[\frac{\beta\sigma(T_m^4 - T_0^4)}{\alpha V_\infty^3}\right]^{q/(\nu+q)} \qquad (8)$$

The reader will note that when $\nu = 0$, corresponding to an atmosphere of uniform temperature and molecular weight, equation (8) reduces to equation (18a*), since $b = q/T_1$. Again the radius of a spherical meteoroid of density, ρ_s is $s = (3m)/(\rho_s B)$.

For a temperature gradient of $+4°$ per kilometer and $V_\infty = 23.1$ km./sec., equation (8) yields a value of m/B some 10% greater than equation (18a*) for an identical atmospheric density and density gradient at the critical velocity in the isothermal case. Since these solutions are in such

close numerical agreement, we may trust equation (8) to give a rather precise result for most problems of micro-meteorites.

2. The Solution in the General Case.—When the physical data justify a more precise theory for micro-meteorites than that given in the previous section, a numerical integration will be required generally. The atmospheric temperature will certainly vary with height in some complex fashion while its mean molecular weight may also vary appreciably. The drag coefficient and the emissivity may depend upon the velocity and temperature of the micro-meteorite and possibly upon the air density. In addition the heat capacity, C_s, may be appreciable and will vary with T_s.

Before going on to the most general situation, let us investigate the error introduced by neglecting the heat capacity in the previous solutions. Beginning with equation (7*), the temperature differential can be related to the velocity differential by means of equation (13*), (14*) and (15*) so that the time, height and atmospheric density are explicitly eliminated. The temperature derivative with respect to the velocity is

$$\frac{dT_s}{dV} = -uV - \frac{w(T_s^4 - T_0^4)}{V^2 \log (V/V_\infty)}, \tag{9}$$

where the constants, u and w, are given by

$$u = \alpha/(C_s D) \tag{10}$$

and

$$w = \beta B\sigma/(bmC_s \cos Z). \tag{11}$$

Our previous assumption in neglecting the heat capacity was derived by equating the right member of equation (9) to zero. The resulting equation (16*) is rewritten:

$$T_s^4 - T_0^4 = -uV^3(\log V/V_\infty)/w. \tag{12}$$

We may now derive a second approximation to the temperature by substituting the derivative of equation (12) into equation (9). This operation leads to the expression

$$T_s^4 - T_0^4 = -\frac{u}{w} V^3 \log (V/V_\infty) \left[1 - \frac{V + 3V \log (V/V_\infty)}{4wT_s^3}\right]. \tag{13}$$

If we now differentiate equation (13) and set $dT_s/dV = 0$ the result will be a second approximation to the velocity-temperature relation at maximum temperature, T_m. The equation is written more simply with the first order auxiliaries x and y where

$$x = V_c/(4wT_m^3), \tag{14}$$

and

$$y = 1 + 3 \log (V_c/V_\infty). \tag{15}$$

After dropping non-zero coefficients, we find the relation at maximum temperature to be

$$y(1 - xy) = x(y - 1)(1 + y/3). \tag{16}$$

If we drop second-order terms in x and y, equation (16) simplifies to

$$y = -x = -\frac{V_\infty e^{-1/3} e^{y/3}}{4wT_m^3} = -\frac{V_\infty e^{-1/3}}{4wT_m^3}. \tag{17}$$

Hence, under the same conditions,

$$\log (V_c/V_\infty) = (y - 1)/3, \tag{18}$$

and

$$V_c^3 = V_\infty^3(1 + y)/e. \tag{19}$$

Thus V_c^3 is reduced by a first-order term ($-y$, which is positive). A substitution of the corrected value of V_c into equation (13) to evaluate T_m, however, leaves no first-order correction terms. The original value of T_m^4 is unchanged except for second and higher-order terms ($1 - 2y^2 + \ldots$). It is clear that no successive approximations can introduce a first-order term into T_m^4. We must conclude, therefore, that the inclusion of the heat-capacity effect into our solution for the maximum temperature of the micro-meteorite reduces by a first-order term the velocity at which the maximum occurs, and correspondingly lowers the atmospheric height and increases the density; but it does not change the maximum temperature by a first-order term.

The value of the correction term, y, expressed in terms of the original physical quantities involving m/B from equation (18a*), may be written

$$y = -3e^{1/3} C_s D T_m (1 - T_0^3/T_m^3)/(4\alpha V_\infty^2). \tag{20}$$

The quantity y is a maximum at the minimum value of V_∞, 11.2 km./sec. We may adopt an extremely high value of $C_s = 0.3$ cal./g. for irons or stones ($C_s = 1.3 \times 10^7$ ergs/g.), $\alpha = 0.9$, $D = 2.2$, $T_m = 1600°$K., and neglect T_0. The resulting value of y appears to be a maximum for any micro-meteorites traversing the earth's atmosphere undamaged. This value is $y = -0.063$. Since the correction term to $T_m^4 - T_0^4$ is $1 - 2y^2, \ldots$, the maximum correction is -0.008, less than 1%.

Of critical importance is the fact that the cos Z term cancels out of y in equation (20). This cancellation means that the correction term is independent of the angle at which the micro-meteorite strikes the atmosphere. Even though relatively massive bodies may be stopped without melting when cos Z is quite small, the effect of heat capacity is no more serious than

for a much smaller body striking vertically. As cos Z decreases, the time available for equilibrium increases in such a fashion as to cancel the correction term.

The question of the mean free path of air molecules compared to the dimensions of the micro-meteorites must be investigated. Since the momentarily trapped air molecules leave the micro-meteorite at the thermal velocity, v_T, of the surface, their density near the surface tends to increase above the ambient air density by a factor of the order of V/v_T°. The impinging air molecules must pass through this haze of escaping gas for a distance comparable to the diameter of the body. If the mean free path of air molecules in the region surrounding the micro-meteorite near the time of maximum surface temperature is greater by a factor of V/v_T than the diameter of the body, we may be reasonably certain that the present theory requires no correction for this effect. There is every reason to believe that the mean free paths of the air molecules at the high velocities here encountered will be greater than those calculated at normal temperature. Hence, if our limit is a safe one for air at normal temperature, it will certainly be safe under the conditions of the present theory.

An inspection of table 1 in the next section shows that the maximum diameter of an undamaged micro-meteorite at vertical incidence velocity is of the order of 0.01 cm. or smaller. The maximum air density at maximum surface temperature is of the order of 6×10^{-9} g./cm.^{3} At normal temperature the mean free paths of air molecules at this density is of the order of 1 cm. Hence the mean free path exceeds the diameter of the micro-meteorite by a factor of approximately one hundred times. The ratio V/v_T is only of the order of ten. Hence the effect of mean free path may be neglected unless cos $Z < 1/10$, in which case $2s$ may be comparable to the standard mean free path. In fact the approximation is rather good for radiants to about 1° of the horizon ($Z = 89°$) or even farther, because of the effective increase of mean free path with velocity and because the effect of molecular encounters changes the theory slowly as the mean free path decreases.

Since ρ_m and s vary roughly in the same fashion with V_∞, and since s varies as sec Z, the problem of mean free path is a function of Z only. For radiants below an altitude of about 1° the calculated maximum dimensions are somewhat underestimated in the present formulae because the accommodation coefficient may be reduced by the shielding effect.

The time-lag in heat transfer from the front to back surface of the micro-meteorite may now be investigated. Suppose that the micro-meteorite does not rotate. A slight development of the "cold wave" approximation of heat transfer as used by L. R. Ingersoll and O. J. Zobel[4] can be applied to this case. A temperature differential over a distance L will be reduced to $1/e$ of its value in a time $\pi L^2/(2H^2)$, where H^2 is the heat diffusivity of the

material. For ordinary stone $H^2 \cong 0.005$ cm.^{2}/sec. For solid metals its value is much larger. With $H^2 \cong 0.005$ and $L = 0.01$ cm. (the largest value of $2s$ for $\cos Z = 1$), the time is 0.03 sec. Unless the meteoric material is a much poorer conductor than ordinary stone, the lag in heat transfer will not be serious for the largest micro-meteorites at vertical incidence. At lower angles of incidence the time of travel increases as sec Z, as does the dimension of the body. Hence the time lag increases as sec Z. The lag begins to introduce appreciable error for $Z > 87°$. The effect will reduce the radiation rate somewhat below its assumed value in the present theory. The calculated critical dimension of a non-rotating micro-meteorite will be overestimated at the lowest velocities in the range $87° < Z < 89°$. If the body is rotating rapidly the error introduced by heat lag will usually be small, even to $Z = 89°$.

The present theory, however, underestimates the critical dimensions of micro-meteorites at extremely large zenith angles because the curvature of the Earth is neglected. In no case can the curvature of the path completely compensate for this effect in view of the fact that the critical velocity is 0.716 V_∞, while the velocity in a circular orbit is 0.707 times the minimum velocity of fall from infinity. We may roughly define the limits set by curvature in the following fashion: while the micro-meteorite is falling through a vertical distance $1/b$, in which the air density increases by a factor e, the change in height introduced by curvature, h, must not exceed δ/b, where δ is small. The horizontal distance, S, then becomes

$$S = (\tan Z)/b. \tag{21}$$

The corresponding change in height, Δh, arising from curvature over the horizontal distance, L, is given to sufficient accuracy by

$$\Delta h = S^2/2R, \tag{22}$$

where R is the radius of the Earth.

If we substitute $\Delta h = \delta/b$ in equation (22) and substitute S from equation (21) we find

$$\tan^2 Z = 2bRS. \tag{23}$$

In an atmosphere of constant temperature gradient, the value of b becomes

$$b = \frac{q + \nu}{T_1}\left(\frac{\rho_1}{\rho}\right)^{\nu/(q+\nu)}. \tag{24}$$

Hence the limiting value of zenith distance, $Z_{max.}$, from equation (23) can be shown to be given by

$$\tan^2 Z_{max.} = \frac{2R(q + \nu)\delta}{T_1}\left(\frac{\rho_1}{\rho}\right)^{\nu/(q+\nu)}. \tag{25}$$

For $\delta = 1/100$, or 1% accuracy, the values of $Z_{max.}$ lie in the range of 70° to 77° for V_∞' having minimum and maximum values, respectively, applying to solar-system micro-meteorites. For $\delta = 1/10$, the corresponding range in $Z_{max.}$ is 84° to 86° (see next section).

As a consequence of the above several considerations we may neglect effects arising from heat capacity, secondary encounter processes and heat transfer in the above explicit theory of undamaged micro-meteorites to limits of zenith angle at which curvature of the earth becomes appreciable. The precision, within the limits of the assumptions, is within 1% for all micro-meteorites striking at Zenith angle less than roughly 70% and within about 10% for zenith angles between 70 and 84%. The theory probably represents the general order of magnitude to a zenith angle of about 87°, when effects of earth curvature become serious for all velocities, while the incomplete heat transfer may introduce further errors at low velocities.

Since Earth curvature and curvature of path can be allowed for in the general theory where numerical integration is used, this theory can be trusted to high precision to a zenith angle of approximately 87° for the minimum velocity. At velocities exceeding 20 km./sec., the theory should be reliable to a zenith angle of about 89° or perhaps a little greater.

The fundamental equations to be solved numerically are:

$$\frac{dV}{dt} = -\frac{AD\rho V^2}{2m}, \qquad (11^*)$$

and

$$T_s^4 - T_0^4 = \frac{\alpha A \rho V^3}{2\beta B \sigma}, \qquad (26)$$

and

$$dt = -dh/(V \cos z), \qquad (13^*)$$

in case Earth curvature may be neglected.

Otherwise h and t must be related by more accurate expressions involving earth curvature and the effect of gravity introduced in the motion. The atmospheric density, ρ, must be adopted as a function of height, h, from some standard atmosphere.

Until a more detailed theory is adopted, such as that by Miss Heineman, the drag coefficient, D, may be determined by the expression

$$D = 2 + \frac{8}{9V}\left(\frac{8k_1 T_s}{\pi \mu_1}\right)^{1/2}. \qquad (27)$$

In the above equations (11*, 26, 13*, 27) only m, T_0, k_1, and σ may be taken as strictly constant in the most general case. The numerical integration, however, still remains a relatively simple type of operation.

The author has met with little success in calculating the degree of surface deterioration that may occur to a micro-meteorite at higher velocities when atmospheric molecules impinge on the surface with energies of several hundred electron volts. Near the lower limit of velocity, 11.2 km./sec., the energies are so small that the deterioration should not be appreciable. Near the upper limit of solar-system velocities, 72 km./sec., the atmospheric molecules must penetrate deeply below the surface, dissociate and damage the lattice structure of the solid. Since the total air mass encountered before the velocity becomes greatly reduced is comparable to the mass of the micro-meteorite, this damage may be considerable. A crude estimate places the critical velocity in the neighborhood of 30 km./sec. Above this velocity the damage probably becomes appreciable and near

TABLE 1

NUMERICAL VALUES FOR MICRO-METEORITES

V_∞, KM./SEC.	D	ρ (MAX. T), G./CM.^{2}	NACA MASS*/ SURFACE, G./CM.^{2} × 104	V-2 MASS*/ SURFACE, G./CM.^{2} × 104
11.3	2.09	5.6×10^{-9}	41.4	24.3
15.0	2.07	2.4×10^{-9}	19.2	11.8
20.0	2.05	1.0×10^{-9}	8.8	5.6
23.1	2.04	6.5×10^{-10}	6.0	3.9
25.0	2.04	5.2×10^{-10}	4.8	3.2
30.0	2.03	3.0×10^{-10}	2.94	1.99
40.0	2.02	1.3×10^{-10}	1.36	0.96
50.0	2.02	6.5×10^{-11}	0.74	0.54
60.0	2.02	3.7×10^{-11}	0.46	0.34
70.0	2.01	2.4×10^{-11}	0.30	0.23

* The tabulated quantity is radius in microns for a spherical body of density $3(s = 3m/B\rho s)$ at vertical incidence. At other angles of incidence the quantities should be multiplied by secant zenith distance of the radiant. See equation (18c*), *et seq.*, for dimensions of non-spherical particles.

70 km./sec. may become quite destructive. At these velocities, however, the maximum dimensions of micro-meteorites are the order of only one micron (table 1, above), rather small for easy study at the earth's surface.

3. Numerical Results.—The atmospheric density, ρ_m, at the maximum surface temperature of the micro-meteorite is given by equation (20*). Although this equation is derived on the assumption of an atmosphere of constant temperature, molecular weight and gravity, nevertheless, it gives the correct order of magnitude for ρ_m. The constants, α, accommodation coefficient and, β, emissivity, we have already adopted as nearly unity and, for lack of better information, equal. If we adopt a maximum temperature, T_m, equal to 1600°K., and $T_0 = 300$°K. then equation (20*) becomes numerically,

$$\rho_m = 8.1 \times 10^9 / V_\infty^3, \tag{28}$$

where c. g. s. units are used.

Numerical values of ρ_m are given in the third column of table 1 for various values of V_∞ in the first column.

The value of the drag coefficient, D, is determined from equation (22*) with the substitution of V_∞ for V. A mean molecular weight of 26.5 and the maximum temperature, T_m, are used in calculating, D. The resulting numerical equation is

$$D = 2 + 1.0 \times 10^5 / V_\infty \tag{29}$$

where V_∞ is expressed in cm./sec.

Numerical values of D are given in the second column of table 1.

The limiting ratio of mass to total surface area, m/B, for micro-meteorites is calculated only from equation (8), in an atmosphere of constant temperature gradient. Vertical incidence is assumed so that $\cos Z = 1$. The remaining constants depend upon the atmospheric conditions. Two solutions are given, one based upon the Tentative Standard Atmosphere of the National Advisory Committee for Aeronautics[5] and the second based upon atmospheric results from the V-2 rocket firings.

The data adopted as a close approximation to the NACA atmosphere are: $h = 0$ at 83 km. altitude, $\rho_1 = 3.32 \times 10^{-8}$ g./cm.^{3}, $T_1 = 240°$K., $\mu = 26.5$, $g = 950$ c. g. s., $q = 3.05 \times 10^{-4}$ °C./cm. and $\nu = 3.65 \times 10^{-5}$ °C./cm. The temperature gradient is, therefore, positive with height at a rate of 3.65 °C./km., beginning at 240°K. at an altitude of 83 km. The resulting numerical form of equation (8) is

$$m/B = 3.27 \times 10^{13} D \sec Z \, V_\infty^{-2.68}. \tag{30a}$$

Numerical results from equation (30a) for $\cos Z = 1$ are given in the fourth column of table 1.

Typical results for atmospheric density by measures from V-2 rockets are given by E. Durand[6] of the Naval Research Laboratory. Adopted are: $h = 0$ at 100 km. altitude, $\rho_1 = 1.00 \times 10^{-9}$ g./cm.^{3}, $T_1 = 223°$, $\mu = 26.5$, $g = 950$ c. g. s., $q = 3.05 \times 10^{-4}$ °C./cm. and $\nu = 5.6 \times 10^{-5}$ °C./cm. Hence the temperature, beginning at 223° at 100 km., increases with height at a rate of 5.6°C./km. The numerical form of equation (30) is

$$m/B = 2.55 \times 10^{12} D \sec Z \, V_\infty^{-2.53}. \tag{30b}$$

This equation leads to the results given in the fifth column of table 1. A comparison of columns four and five shows that the V-2 atmosphere leads to smaller values of m/B than the NACA atmosphere in a ratio of about 2/3, little dependent upon velocity. This difference arises largely from the difference in logarithmic density gradient between the two atmospheres

at corresponding values of the density. The critical heights range from just below 100 km. to roughly 140 km.

The calculations of the limits of zenith distance at which the curvature of the Earth introduces appreciable error, given by equation (25) and presented near the end of the previous section, were derived from the V-2 atmosphere. This atmosphere gives less favorable limits than the NACA atmosphere.

Quantities in table 1 are also presented specifically for $V_\infty = 23.1$ km./sec., corresponding to the Giacobinid meteor shower. In the Washington, D. C. area, cos Z was about 0.45 for this shower at the time of maximum activity. Hence for undamaged iron micro-meteorites of density 7.7, the maximum spherical radii should have been 5.2 microns for the NACA atmosphere and 3.4 microns for the V-2 atmosphere. For stony-irons of density 4.0 these values become 10.0 and 6.5 microns, respectively. For long cylinders of circular cross-section, the minor diameter is 4/3 the spherical radius, giving a predicted range from 4.2 to 13.3 microns, the range depending upon uncertainties in the atmosphere and in the density of the micro-meteorites.

Coincident with the shower, allowing for time of fall, H. E. Landsberg[7] found at Mt. Weather, Va., "wedge-shaped opaque" magnetic particles of dimensions 40×5, 40×5 and 100×4 microns. These he attributed to the shower. The agreement of the minor dimension with the present predictions is entirely satisfactory. He also found 3 "round" particles of diameter 5 to 10 microns, again in agreement with this theory. On the other hand, he found a few "round" particles of diameters from 20 to 40 microns. Although these are too large for the present theory, there is no reason to doubt that larger bodies may have been melted and stopped with minor surface vaporization, which would certainly tend to smooth the surface—in keeping with Landsberg's observations. Opik[8] discusses the formation of droplets in meteoroids in considerable detail.

Landsberg estimates roughly a rate of fall of $3^1/_2$ miles per day for the particles of diameter 4 to 10 microns, 13 miles per day for those of diameter 20–30 microns and 100 miles per day for the one of diameter of 60 microns. The expected times of fall would be quite variable for the smaller particles because of atmospheric turbulence.

Generally, then, the present theory is in good agreement with Landsberg's observations. The critical dimensions here calculated are of most value in predicting a dimension that divides irregular rough-surfaced particles from larger ones that have been melted and partially vaporized in passing through the atmosphere.

It is now clear that the dependence of critical particle size on assumed atmospheric conditions is too small to make the study of micro-meteorites valuable at this time as a measure of atmospheric density gradient. Nevertheless, the minor dimensions of the largest rough irregular wedge-shaped

pieces observed by Landsberg fall closer to the prediction of the V-2 atmosphere than the NACA atmosphere. It is difficult to know how to weight the significance of these results, particularly in view of uncertainties in the physical characteristics of micro-meteorites.

Generally the present theory may be most useful in the following respects:

(*a*) It provides theoretical evidence that micro-meteorites can fall undamaged, hence gives impetus to their collection and study, which should lead to measures of space densities for small particles.

(*b*) It can aid in distinguishing real micro-meteorites from terrestrial objects.

(*c*) It can aid in identifying micro-meteorites from two nearly coincident meteor showers.

(*d*) It can assist in determining the original velocity of recognized micro-meteorites.

(*e*) It may help provide evidence concerning the final dimensions of particles producing meteors.

The collection and laboratory study of micro-meteorites from meteoric streams can be invaluable in establishing the nature of interplanetary material, specifically cometary material. If the writer's present working hypothesis is correct, viz., that comets were formed at great solar distances by the condensation of gas at very low temperatures, then micro-meteorites from meteor showers may represent the only laboratory samples of typical or atypical interstellar solids. Ordinary meteorites appear to arise from a different source.[9]

On a purely speculative level is the possibility of detecting evidence for past solar-system catastrophies. If a minor or major planetary disruption actually occurred in recent astronomical time, as suggested by C. A. Bauer[10] to account for the meteoritic helium contents, evidence for the concomitant micro-meteorites may conceivably be found in the chalk beds of the Cretaceous Period or in other geological formations.

Furthermore, there is the possibility that micro-meteorites in deep oceanic sediments (as suggested by the *Challenger Expedition*), in glacial snows or even in geological formations (such as chalk beds) may provide a record of past astronomical activities.

The writer earnestly hopes that this paper may encourage the collection and study of micro-meteorites. Current work on this problem is being conducted by D. K. Norris and F. S. Hogg,[11] while J. D. Buddhue[12] has recently summarized his and others' observation data on meteoritic dust in general.

* This paper was written as a part of the investigation for United States Naval Ordnance Contract-10449-05512.

[1] Whipple, F. L., these PROCEEDINGS **36**, 687–695, 1950.

[2] See, e.g., Mitra, S. K., *The Upper Atmosphere*, ch. III, The Royal Asiatic Society of Bengal, 1948.

[3] See, e.g., Mitra, S. K., *loc. cit.*, p. 6.

[4] *Mathematical Theory of Heat Conduction*, Ginn and Co., 1913, p. 40.

[5] Warfield, C. N., NACA, Tech. Note, No. 1200, Langley Field, Jan., 1947.

[6] *The Atmospheres of the Earth & Planets*, Univ. of Chicago Press, 1948, p. 142.

[7] *Pop. Ast.*, **55**, 322 (1947).

[8] *Pub. Univ. Tartu*, **29**, No. 5, 51 (1937).

[9] See, e.g., Brown, H., and Patterson, C., *J. Geol.*, **56**, 85 (1948).

[10] *Phys. Rev.*, **74**, 501 (1948).

[11] Reported at the June 1949 meeting of the Am. Ast. Soc.

[12] Univ. N. Mex. Pub. in Meteoritics, No. 2, 1950.

*MULTIMOLECULAR ADSORPTION OF HORSE HEART METMYOGLOBIN ONTO FILMS OF BARIUM STEARATE**

BY ALBERT A. FISK†

GATES AND CRELLIN LABORATORIES OF CHEMISTRY, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA 4, CALIFORNIA‡

Communicated by Linus Pauling, October 31, 1950

In a previous paper[1] it was shown that some protein solutions, when adsorbed onto films of barium stearate, followed the classical Langmuir adsorption isotherm. The thicknesses of these unimolecular layers were calculated by using the Langmuir equation, and it was shown that these values for human and horse hemoglobin and for bovine serum albumin agreed with values in the literature for the thicknesses of these molecules.

It was felt advisable to test this technique on another protein of size and molecular weight different from the above proteins. Myoglobin was chosen because its dimensions have been recently determined by x-ray diffraction.[2]

Experimental.—The technique was adequately described in the previous paper. The thicknesses of the metmyoglobin layers adsorbed onto metallic slides covered with an optical gauge of barium stearate were determined on the ellipsometer, an optical instrument which measures the ellipticity of polarized light reflected from a metallic surface.

The metmyoglobin solutions were prepared from fresh horse heart by repeated precipitations from strong phosphate buffers. The Beckman spectrophotometer measurement of a carbonmonoxylated solution did not show the presence of any contaminating hemoglobin. The electrophoresis of a metmyoglobin solution showed that two components were present, the smaller component representing 10% of the total area. An electrophoresis of a carbonmonoxymyoglobin solution showed that the smaller component was present in one-half its concentration in the metmyoglobin

preparation. A careful separation by electrophoresis was made on a metmyoglobin solution, but subsequent electrophoreses of the separated main component revealed that the same smaller component was still present in 10% concentration. It was, therefore, assumed that the smaller component was probably not an impurity, but was metmyoglobin in the dimer or higher form. The adsorption experiments were done on this fraction at pH 6.8 in phosphate buffer of ionic strength 0.1.

Results.—The results may be seen in figure 1. The thicknesses in figure 1 are not equivalent barium stearate thicknesses, as in the previous paper, for refractive index corrections due to the relative humidity of the laboratory and the assumed method of packing of the adsorbed myoglobin

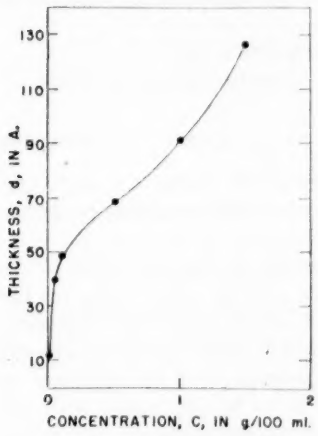

FIGURE I
ADSORPTION ISOTHERM FOR HORSE HEART METMYOGLOBIN ADSORBED ONTO BARIUM STEARATE
t = 16-19° C

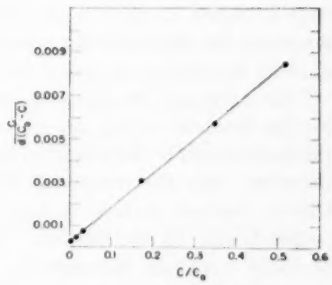

FIGURE II
B.E.T. PLOT FOR HORSE HEART METMYOGLOBIN ADSORBED ONTO BARIUM STEARATE
C_0 = 2.90 g/100 ml.

molecules led to small increases in the thicknesses. It was assumed, on the basis of x-ray diffraction data[2] and the preliminary results, that the packing model would be best approximated by assuming the closest packing of elliptical cylinders lying on their sides, the voids being filled with air. The thickness corrections were calculated by using the derivative of the Drude equation,[3]

$$\frac{dl}{l} = \frac{2dn}{n(n^2 - 1)},$$

where

l = equivalent thickness of barium stearate,

n = refractive index of barium stearate.

Interpretation of Results.—The results as shown in figure 1 indicate multimolecular adsorption of metmyoglobin onto barium stearate. The Brunauer, Emmett and Teller equation[4] for multimolecular adsorption is

$$\frac{C}{d(C_0 - C)} = \frac{1}{d_m b} + \left(\frac{b - 1}{d_m b}\right)\frac{C}{C_0},$$

where

C = concentration of metmyoglobin solution,
d = calculated thickness,
C_0 = constant, the concentration at infinite thickness, assumed to be 2.90 g./100 ml.,
b = constant related to the heat of adsorption, and
d_m = thickness of a unimolecular layer of metmyoglobin molecules.

It is seen in figure 2 that a plot of $\frac{C}{d(C_0 - C)}$ against C/C_0, using the data in figure 1, gives a straight line. From the calculated slope and intercept, the thickness of a unimolecular layer of metmyoglobin molecules, d_m, was determined to be 62 A.

This value is to be compared with the maximum probable value, 57 A., for the diameter of the myoglobin molecule and the value, 9 A., for the thickness found by Kendrew[2] in his x-ray diffraction study. It would seem, therefore, that the metmyoglobin molecule has been adsorbed onto the barium stearate surface so that a long dimension, not the thickness, is perpendicular to the surface. This case is to be contrasted with the previous work[1] on hemoglobins and bovine serum albumin, where the short dimension was perpendicular to the surface. However, since myoglobin is known to have but one heme and since this heme is presumed to be attached to the side of the molecule,[2] it is not surprising to find that myoglobin may be adsorbed with the heme in a preferred orientation, either against the barium stearate surface or away from it.

The multimolecular adsorption of myoglobin, in contrast to the unimolecular adsorption of the other protein molecules studied, is not entirely unexpected in view of the fact that Taylor[5] found a considerable amount of heme-heme interaction in magnetic susceptibility measurements of myoglobin solutions. The ability of myoglobin molecules to associate was also apparent in the aforementioned preparative electrophoreses.

From the above experiment, it is apparent that the determination of protein dimensions from measurements of thin films of proteins adsorbed onto a barium stearate surface is not limited to molecules in the hemoglobin-albumin class. However, it is equally apparent that the interpreta-

tion of the results may depend on other information concerning the shape and dimensions of the molecule.

Summary.—The thickness of a unimolecular layer of horse heart metmyoglobin molecules adsorbed onto barium stearate was calculated to be 62 A. by means of the B.E.T. equation for multimolecular adsorption.

* This work was supported in part by grants from the American Cancer Society and the U. S. Public Health Service.

†American Cancer Society Fellow. Present address: Biochemical Laboratory, University of Virginia Medical School, Charlottesville, Va.

‡ Contribution No. 1486.

[1] Fisk, A. A., PROC. NATL. ACAD. SCI., **36,** 518 (1950).

[2] Kendrew, J. C., in "Haemoglobin," ed. by Roughton, F. J. W., and Kendrew, J. C., Interscience Publishers Inc., New York, N. Y., 1949, page 149.

[3] Drude, P., *Ann. d. Physik. und Chemie,* **36,** 865 (1889).

[4] Brunauer, S., "Physical Adsorption," Princeton University Press, Princeton, N. J., 1943, Ch. VI.

[5] Taylor, D. S., *J. Am. Chem. Soc.,* **61,** 2150 (1939).

A COMPARISON OF THE CONTENT OF DESOXYRIBOSENUCLEIC ACID (DNA) IN ISOLATED ANIMAL NUCLEI BY CYTOCHEMICAL AND CHEMICAL METHODS

BY CECILIE LEUCHTENBERGER, ROGER VENDRELY* AND COLETTE VENDRELY*

INSTITUTE OF PATHOLOGY, WESTERN RESERVE UNIVERSITY, CLEVELAND, OHIO

Communicated by C. W. Metz, November 7, 1950

In 1948, Boivin, Vendrely and Vendrely[1, 2] presented evidence that the amount of desoxyribosenucleic acid (DNA) in nuclei of different tissues was constant for the same animal and twice that of the sperm. These authors suggested that such a constant relationship applies to all diploid animal cells and might be an expression of the genetical equipment of the cells.

While some recent work[3–6] lends support to this idea, other studies stand in contradiction. In 1949, Mirsky and Ris[6] reported analyses on the content of DNA in nuclei of various tissues of *mammals* which showed a quite different relationship to the sperm from that reported by Boivin, Vendrely and Vendrely. According to Mirsky and Ris, in calf thymus, calf lymph nodes, beef kidney and beef liver, the amount of DNA in diploid nuclei in relation to the amount of DNA in the sperm varied from 2.5 to 3:1, in contrast to the 2:1 ratio found for the same nuclei by Boivin, Vendrely and Vendrely.

In 1950, Pasteels and Lison,[7] working on the rat, reported that the DNA in nuclei of diploid liver and pancreas cells showed a relationship to the sperm which was only 1.3–1.5:1, which is again a significant departure from the ratio found by the French workers.

Considering that Mirsky and Ris found values of DNA for *mammalian* diploid nuclei which are much too high while Pasteels and Lison found values which are much too low in relation to the sperm, such deviations, if confirmed, would, of course, be a strong argument against the concept of Boivin, Vendrely and Vendrely that the amount of DNA in all diploid somatic nuclei of animals is exactly twice that of the sperm.

A reinvestigation of this particular problem therefore seemed essential and it was decided to make a determination of the DNA on the same preparations of isolated animal nuclei by two independent, quite different methods: the chemical method, in which the amount of DNA per nucleus is computed from a large number of nuclei subjected to biochemical analysis, and the cytochemical method, which allows a direct estimation of DNA in a single nucleus through photometric measurements. In the present paper we report the results on the amount of DNA in nuclei of these two methods when applied to various tissues of beef and rat.

Material and Methods.—Liver, kidney, spleen and sperms were obtained from freshly killed beef and rats, the nuclei isolated and the amount of DNA determined chemically as previously described by Vendrely and Vendrely.[2] For the cytochemical analysis, the same suspension of isolated nuclei was used for smears, which were fixed in Carnoy and 50% Formalin and on which the Feulgen reaction was carried out. The DNA was estimated by photometric measurements with an apparatus similar to that of Pollister and Moses,[8] and by the method as described by Schrader and Leuchtenberger.[4] The cytochemical analysis was also done on sections of the same organs fixed in Carnoy and 50% Formalin.

Results.—In the following table the results of the cytochemical and chemical analysis of DNA in isolated nuclei of different tissues of beef and rat are presented. It should be pointed out that the amount of DNA per nucleus found by the cytochemical analysis is a mean value expressed in arbitrary units (as previously described),[4] while the amount of DNA per nucleus by the chemical analysis is given in absolute amounts in 10^{-6} γ.

If we consider first the cytochemical analysis of the DNA of different tissues of the beef, it is evident that the liver, spleen and kidney contain approximately the same amount of DNA per nucleus and twice that of the sperm. The proportion is also maintained if we look at the data obtained by the chemical analysis on the same suspension of nuclei. If we now examine the cytochemical analysis of the DNA in the tissues of the rat, it is evident that in the liver the nuclei fall into three classes with values of DNA per nucleus of 3.3, 6.6 and 12.1 (arbitrary units), respectively.

Furthermore, these values represent nearly exact multiples of each other. The picture is quite different if we consider the amount of DNA of the nuclei of the kidney of the rat. Here only a mean value of 3.3 per nucleus was observed, which corresponds to the lowest class of DNA found in liver nuclei.

In the chemical analysis, all the size classes of nuclei in the liver are, of course, thrown together and the results of such lumping is expressed by the fact that the amount of DNA per liver nucleus is considerably higher than that of the kidney.

Discussion of the Results.—In evaluating the data it should be pointed out that the cytochemical values given in the table for the amount of DNA per nucleus of a tissue actually represent a mean of a number of individual DNA values. While this mean amount is constant and the same for diploid nuclei of different tissues for the same animal there is a variation from nucleus to nucleus within each tissue, sometimes as high as 50%. For example, the beef spleen which gave an average amount of DNA of 3 per nucleus, similar to the average amount for beef kidney and beef liver, showed individual nuclei with values of DNA ranging from 2.7 to 4.0. Whether such differences are due to errors inherent in the cytochemical photometric method as previously discussed,[9] or are due to biological variations in the amount of DNA among individual nuclei of a tissue cannot be decided at the present time.

TABLE 1

COMPARISON OF THE AMOUNT OF DESOXYRIBOSENUCLEIC ACID (DNA) IN ISOLATED NUCLEI OF DIFFERENT TISSUES OF BEEF AND RAT BY CHEMICAL AND CYTOCHEMICAL ANALYSIS

TYPE OF ANIMAL	TYPE OF TISSUE	CYTOCHEMICAL ANALYSIS: NO. OF NUCLEI MEASURED	CYTOCHEMICAL ANALYSIS: MEAN AMOUNT OF DNA PER NUCLEUS (ARBITRARY UNITS)	CHEMICAL ANALYSIS, AMOUNT OF DNA PER NUCLEUS (10^{-6} γ)
Beef	Liver	30	3.3 ± 0.11	6.8
Beef	Spleen	23	3.0 ± 0.08	7.0
Beef	Kidney	26	3.3 ± 0.10	6.4
Beef	Sperm	25	1.6 ± 0.03	3.3[a]
Rat	Liver	15	3.3 ± 0.08	
Rat	Liver	21	6.6 ± 0.24	8.2
Rat	Liver	7	12.1 ± 0.50	
Rat	Kidney	38	3.3 ± 0.06	5.5

[a] This value is taken from previous analyses by Vendrely and Vendrely.[3]

NOTE: The amount of DNA determined by cytochemical analysis is given in arbitrary units.

If we compare now the results obtained by the cytochemical and the chemical analyses of DNA in the nuclei of different tissues of the beef, it is evident that by both analyses the diploid nuclei of the somatic tissues show twice the amount of DNA of the sperm. Therefore, the unavoidable

conclusion follows that constancy of the DNA in nuclei of the beef, as first reported by Boivin, Vendrely and Vendrely in 1948, holds true, even when individual nuclei are analyzed for their content of DNA (as done by the cytochemical studies), instead of analyzing a large mass of nuclei and computing the value for one nucleus. These results, therefore, do not confirm the measurements of DNA of *mammalian* nuclei as reported by Mirsky and Ris.[6] According to the last-named authors, the higher values of DNA in the beef liver (which represent three times the value of the sperm—namely, 8.4 to 2.82) cannot be explained by the presence of polyploid nuclei, for they state: "Polyploidy is not the explanation of the high values found for beef liver, for in counting liver nuclei the double sized nuclei were counted separately and the necessary correction was made."

In view of the fact that the results which we obtained in the present study by both cytochemical and chemical analyses are in complete agreement with the findings previously obtained on beef by Vendrely and Vendrely, the discordant findings of Mirsky and Ris remain unexplained. It might be of interest to state that in the cytochemical analysis of DNA of the nuclei of the beef liver only two nuclei among thirty-two examined showed amounts of DNA which were twice that of the values reported. A similar relationship holds true for spleen and kidney. It should be pointed out, however, that occasionally nuclei were found which had a considerably lower value than 3.3, especially if nuclei of very small size with a pycnotic appearance were measured. Whether these are degenerating, dying nuclei so familiar to every pathologist, or whether the lower value is due to errors in the photometric measurements cannot be decided at this moment with certainty. We are inclined to believe that even in the normal organs there are always a small number of degenerating nuclei present which gradually lose their DNA as demonstrated for pycnotic nuclei by Leuchtenberger.[9]

Turning now to data on rat nuclei, the finding of three distinct classes with multiple values of DNA, 3.3, 6.6 and 12.1 respectively in the analysis of isolated liver nuclei is in good agreement with the three classes of DNA values found by Ris and Mirsky[10] in sectioned material of rat liver. A similar agreement between amounts of DNA found by the cytochemical analysis in *isolated nuclei* and tissue *sections* has been found also in our own studies.

Pasteels and Lison[7] also report three classes of DNA values in the rat liver nuclei in sections but each of these classes has a value which is about 30% too low, if compared with the results in other organs. While kidney, for instance, has one class of DNA with a value of 227, the liver has the three size classes—156, 347, 772. Hence the measurements in the kidney find no equivalent in the liver nuclei where the minimum value is significantly lower, while the two higher values stand in no multiple relation to the single class represented by the kidney nuclei.

We cannot confirm these results because, although we found also only one class of DNA in kidney in contrast to the three classes in the liver as did Pasteels and Lison, the amount of DNA in the nuclei of the kidney of the rat is approximately the same as the lowest amount of DNA in the liver nuclei.

If we now compare the results obtained by our cytochemical analysis of DNA in nuclei of rat liver and rat kidney with the chemical analysis on the same material, it is evident that the higher value of DNA in the liver nuclei—$8.2 \times 10^{-6}\gamma$ as compared with the lower value of $5.5 \times 10^{-6}\gamma$ for the nuclei of the kidney—is an expression of the polyploidy in the liver. We feel that these studies demonstrate rather well the importance of the simultaneous application of the cytochemical and chemical method for the analysis of DNA on the same material. Only by the cytochemical method was it possible to recognize in the rat liver three different types of nuclei which were characterized by three different amounts of DNA (multiples of each other). In contrast the result of the chemical method represents an average value of three types of nuclei with different amounts of DNA and is, therefore, not the true quantity of DNA of the individual nuclei in the liver. This does not by any means lessen the importance of the chemical analysis of DNA in diploid nuclei, particularly if we keep in mind that if the chemist presents a value of DNA for a nucleus it is based on much larger samples of nuclei (billions), while the cytochemist—even if he were able to study 1000 nuclei of one tissue—always analyzes only at comparatively small population of nuclei in respect to the whole tissue. There seems to be no doubt as to the desirability of employing both methods simultaneously on the same material, especially in cases involving polyploidy or in tissues which show pathological changes.

Summary and Conclusions.—DNA was determined in the same preparation of isolated animal nuclei by a cytochemical and a chemical method. Contrary to the findings of Mirsky and Ris, we found by both methods that diploid nuclei of various tissues of beef contain the same amount of DNA and twice that of the sperm. The constancy in the amount of DNA was also observed in the diploid class of liver and kidney nuclei of the rat in contrast to the findings of Pasteels and Lison.

This analysis, based on two different and independent methods, thus gives firm support to the concept of Boivin, Vendrely and Vendrely, that mammalian diploid nuclei of various tissues contain a constant and characteristic amount of DNA and that this amount is double that which is carried by the sperm.[11]

* Faculté de Medecine, Institut de Bacteriologie, Strasbourg, France.

[1] Boivin, A., Vendrely, R., and Vendrely, C., *Compt. rend. Acad. Sci.*, **226**, 1061–1062 (1948).

[2] Vendrely, R., and Vendrely, C., *Experientia*, IV (10), 434–436 (1948).

[7] Davidson, J. N., Leslie, I., Smellie, R. M. S. and Thompson, R. Y., *Biochem. J.*, XL, 46 (1950).
[4] Schrader, F., and Leuchtenberger, C., *J. Exp. Cell Res.*, **1**, 421–452 (1950).
[5] Swift, H. H., *Physiol. Zool.*, **23**, 169–200 (1950).
[6] Mirsky, A. E., and Ris, H., *Nature*, **163**, 666–667 (1949).
[7] Pasteels, J., and Lison, L., *Compt. rend. Acad. Sci.*, **230**, 780–782 (1950).
[8] Pollister, A. W., and Moses, M. J., *J. Gen. Physiol.*, **32**, 567–577 (1949).
[9] Leuchtenberger, C., *Chromosoma*, **3**, 449–473 (1950).
[10] Ris, H., and Mirsky, A. E., *J. Gen. Physiol.*, **33**, 125–146 (1949).
[11] We wish to express our sincere appreciation to Dr. Alan R. Moritz, Director of the Institute of Pathology, Western Reserve University, who has made this combined study possible.

STRUCTURES IN E. COLI RESEMBLING CHROMONEMATA

By Alfred Marshak

New York University, College of Medicine and the Division of Tuberculosis, U. S. Public Health Service

Communicated by Barbara McClintock, November 11, 1950

Although bodies having the characteristics of nuclei have been described in studies of E. coli with the light microscope and with the electron microscope, the internal structure of the nuclei or of the chromosomes has not been observed hitherto.[1–4] The observations reported here show the presence of structures resembling chromonemata, and chemical analyses of the particulate fraction in which the apparent nuclear bodies are concentrated fulfill expectations from previous studies.[3, 4]

Bacteria in the early log phase of growth were suspended in 5% citric acid and immediately subjected to sonic vibration (9 kc.)[5] for 45 minutes and then subjected to differential centrifugation. A heavy black pigmented fraction was obtained at 900 g. and a light unpigmented one at 3600 g. Very little additional solid material was collected at 55,000 g.[6] The average size of the heavy particles was $0.96\ t\ 0.002\ \mu \times 0.63\ t\ 0.002\ \mu$, which was approximately the size of the Feulgen-positive bodies of the unbroken cells.

The heavy fraction was suspended in distilled water and a drop of the suspension on a collodion film supported by wire mesh was frozen on a block of solid CO_2. The frozen material was then evacuated and shadowed with chromium. When examined in the electron microscope many of the particles showed rather regular peripheral lobulations, suggesting the presence of a banded or helical structure. In many instances, helical structures whose coils corresponded with the peripheral lobulations could be observed (figure 1). Most of the particles observed gave some indication of the presence of helices although these structures appeared much

more clearly in some than in others. In some particles, a system of at least two helices could be seen and the constituent strands of each helix were observed to be themselves helically coiled (figure 2). Occasional particles were found whose widths were approximately one-half those of the larger

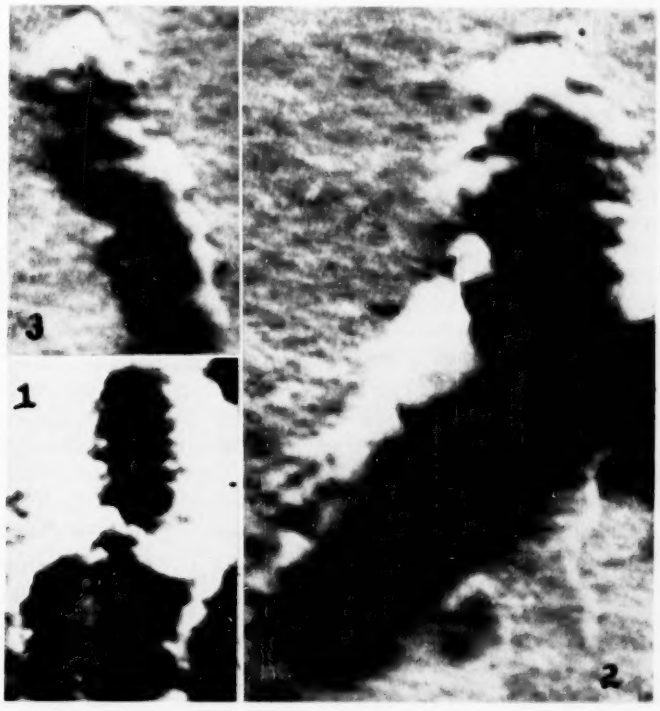

FIGURE 1

Particles from the heavy fraction showing peripheral lobulation and coils of helices. Chromium 17,500 ×.

FIGURE 2

Helical coiling of individual strands composing the larger helices. Chromium 52,500 ×.

FIGURE 3

Particle of one-half the usual width showing the helix composed of 2 strands each of which is itself helically coiled. Chromium 52,000 ×.

particles (figure 3). They could be seen to contain a rather regular helix, the gyres of which were made up of two strands. Where the individual strands could be clearly distinguished, they could also be seen to be helically coiled, with a pitch approximately $^1/_3$ that of the larger helix.

After extraction with cold trichloracetic acid and with fat solvents the intact bacteria and the particulate fractions were digested with perchloric acid and the purine and pyrimidine content determined by chromatography and spectrophotometry.[7, 8] The results obtained are given below:

NUCLEIC ACID BASES (μG./MG. DRY MATTER)

	ADENINE	GUANINE	CYTOSINE	URACIL	THYMINE
Intact bacteria	18.6	24.3	13.9	7.39	3.32
Heavy particles	10.6	15.3	12.1	4.46	2.34
Light particles	19.9	25.1	14.8	8.76	3.85

Each value in the table represents a single determination and is subject to an error of approximately 10%. Absorption spectra were taken of each recovered base, and comparison with the curves for the pure bases showed good agreement except in the case of thymine. The curve for thymine from the intact bacteria gave a fair approximation to the curve for pure thymine, but the curves for material recovered from the particulate fractions deviated markedly.

The contamination has the effect of increasing the apparent concentration of thymine but the extent of contamination cannot be determined with accuracy. The thymine values must therefore be considered to be unreliable. The reasons for the difference in chromatographic behavior of the digests from the whole bacteria and the particles are unknown.

The concentration of each base other than thymine is lower in the heavy particle fraction than in either the light particles or the intact bacteria. The sum of the bases is lower in the heavy particle fraction than in either the intact bacteria or the light particle fraction by 34–35%.

Discussion.—Previous investigations have shown that the cytoplasm of E. coli is more dense than the nucleus,[3] and that the material responsible for the greater density can be removed by procedures for extracting ribonucleic acid.[4] If it is postulated that the heavy particle fraction represents a concentration of nuclear material as suggested by the dimensions of the particles and their internal morphology, a lower concentration of nitrogenous base would be expected. This expectation was fulfilled. A greater proportion of thymine to the total base content would also be expected. However as indicated previously the thymine data are unreliable. As far as they go the chemical data support the hypothesis that there is a concentration of nuclear material in the heavy particle fraction.

The chromatinic bodies described by Hillier, Mudd and Smith[3] are frequently bipartite or bifurcate. Rough measurements from their illustrations give an approximate value of 0.3 μ for the width of a single limb of these bodies. This may be compared with the width 0.29 μ obtained from measurement of particles containing a single large helix (figure 3) and the width 0.63 μ for particles containing two such helices (figures 1 and 2).

If, as has been previously suggested, each limb of the bipartite body be assumed to be a chromosome, then the above calculations suggest that the helical structures may be chromonemata. The bipartite nature of the helix and the presence of "secondary coiling" of the constituent helices are exactly analogous to conditions found in chromosomes of higher plants.[9, 10] On the basis of both the morphological and the chemical evidence therefore it is concluded that the coiled structures here described represent chromonemata of the chromosomes of E. coli.

Coiled structures have been found in biological materials which are not related to chromosomes. The double helix of elastic tissue differs morphologically from the E. coli structures in that the pair of filaments is not itself helically coiled, and also chemically since the coiled fibers of elastic tissue disintegrate at a pH lower than 5.7.[11] The helices of the sperm neck-piece and middle-piece show no "secondary coiling," the latter occurring in groups of 3. Both must be impregnated with chromic or osmic acids to be detected, suggesting a high lipid content.[12]

Summary.—E. coli was fragmented by sonic vibration in the presence of citric acid and the fragments separated by centrifugation into a light and a heavy fraction. Analysis of purine and pyrimidine content showed the latter to represent a concentration of nuclear material. The dimensions of the fragments were comparable to those of the chromatinic bodies of the same cells. Particles in this fraction showed dual helically coiled structures resembling chromonemata of higher plants and differing in morphology from coiled structures of non-chromosomal materials previously described.

[1] Epstein, G. W., Ravich-Berger, E. D., and Svinkina, A. A., *Giorn. Bateriol. Immunol.*, **16**, 1 (1936).

[2] Robinow, C. F., *Proc. Roy. Soc. London*, **130 B**, 299 (1942).

[3] Hillier, J., Mudd, S., and Smith, A. G., *J. Bacteriol.*, **57**, 319 (1949).

[4] Mudd, S., and Smith, A. G., *Ibid.*, **59**, 561 (1950).

[5] Raytheon magnetostriction oscillator (Model S 102).

[6] Servall-SS-2 vacuum angle centrifuge.

[7] Marshak, A., and Vogel, H., *Federation Proc.*, **9**, 85 (1949).

[8] Marshak, A., and Vogel, H., *J. Biol. Chem.* (1950), in press.

[9] Marshak, A., *J. Heredity*, **27**, 459 (1936).

[10] Kaufman, B. P., *Bot. Rev.*, **14**, 57 (1948).

[11] Gross, Jerome, *J. Exp. Med.*, **89**, 699 (1949).

[12] Bretschneider, L. H., *Proc. Konink. Akad. Wetensh. (Amsterdam)*, **52**, 526 (1949).

PARTIAL DIFFERENTIAL EQUATIONS AND GENERALIZED ANALYTIC FUNCTIONS. SECOND NOTE

BY LIPMAN BERS

SYRACUSE UNIVERSITY AND THE INSTITUTE FOR ADVANCED STUDY

Communicated by Marston Morse, November 4, 1950

1. Introduction.—This note (an abstract of results to be published in full elsewhere) contains a reformulation and extension of the theory outlined in the preceding note under the same title[1] (henceforth referred to as I). The new theory is consequently based on a complex differentiation process (and on the concomitant integration) and is therefore a true generalization of classical function theory. It applies to all elliptic linear homogeneous partial differential equations of second order (and to systems of first order equations) for unknown functions of two variables. The new theory contains as special cases and unifies the results of I on self-adjoint equations, the earlier more special theory of sigma-monogenic functions by A. Gelbart and the author,[2] and the very ingenious extension of this theory proposed by Markushevitch.[3]

Notations: $\mathbf{E}_0$ denotes the finite z-plane: $|z| = |x + iy| < +\infty$, $\mathbf{E}$ the whole function theoretical plane (Riemann sphere): $|z| \leq +\infty$ $\bar{\mathbf{D}}$ the closure of a domain $\mathbf{D}$. Functions of the real variables x and y are written as functions of $z = x + iy$, without implying analyticity. If f is a complex number, $\bar{f}$ denotes its complex conjugate. Partial derivatives are denoted by subscripts; in particular, $2w_{\bar{z}} = w_x - iw_y$, $2w_z = w_x + iw_y$. A (complex-valued) function $w(z)$ belongs to class $\mathfrak{C}$ (class $\mathfrak{C}$) if it is continuous (satisfies a Hölder condition). If w has partial derivatives with respect to x and y of order n, which are of class $\mathfrak{C}$ (of class $\mathfrak{C}$), then w is of class $\mathfrak{C}^n$ (of class $\mathfrak{C}^n$). $w(z)$ is said to have property $\mathcal{P}$ at $z = \infty$ if $w(1/z)$ has property $\mathcal{P}$ at $z = 0$.

2. Differentiation and Integration. Pseudo-analytic Functions.—Two functions, $F(z)$ and $G(z)$, of class $\mathfrak{C}^1$ defined in a domain $\mathbf{D}_0 \subset \mathbf{E}_0$ are said to form a *generating pair* in $\mathbf{D}_0$ if $\operatorname{Im}(\bar{F}G) > 0$. Our generalization of function theory is based on assigning to a generating pair (F, G) the part played in the classical theory by $(1, i)$.

A function $w(z)$ defined in a domain $\mathbf{D}$, $\bar{\mathbf{D}} \subset \mathbf{D}_0$, admits the unique representation

$$w(z) = \varphi(z)F(z) + \psi(z)G(z), \tag{1}$$

where φ and ψ are real-valued. If for some $z_0 \in \mathbf{D}$ the limit

$$w(z_0) = \lim_{z \to z_0} \frac{w(z) - \varphi(z_0)F(z) - \psi(z_0)G(z)}{z - z_0} \tag{2}$$

exists and is finite, $w(z)$ is said to possess at z_0 the (F, G)-*derivative* $\dot{w} = d_{(F,G)}w/dz$. It turns out the existence of $\dot{w}$ in a domain has as far-reaching consequences as that of the classical derivative.

A function possessing an (F, G)-derivative at all points of $\mathbf{D}$ is called regular (F, G)-*pseudo-analytic of the first kind* (or, more briefly, pseudo-analytic) in $\mathbf{D}$. If (1) is such a function, we call $\omega(z) = \varphi(z) + i\psi(z)$ *pseudo-analytic of the second kind*. The correspondence between w and ω is one-to-one, and the two functions should be thought of as two representations of the same mathematical entity. The class of pseudo-analytic functions is closed under addition, multiplication by real constants, and uniform convergence.

THEOREM. *If $w(z)$ is pseudo-analytic of the first kind, $w(z)$ and $\dot{w}(z)$ are of class $\mathfrak{C}^1$. If $\omega(z)$ is pseudo-analytic of the second kind, $\omega(z)$ is of class $\mathfrak{C}^2$, and the mapping $\omega = \omega(z)$ is an interior, sense-preserving, quasi-conformal transformation.*

An isolated singularity z_0 of a single-valued pseudo-analytic function $w(z)$ is called a pole if $w(z) \sim a(z - z_0)^{-n}$, $z \to z_0$, where $a \neq 0$ and n is a positive integer, an essential singularity if $w(z)$ comes arbitrarily close to every complex number as $z \to z_0$.

THEOREM. *An isolated singularity of a single-valued pseudo-analytic function is either essential, or removable, or a pole.*

Let $W(z)$ be a continuous function, and Γ a rectifiable curve leading from z_0 to z_1. The (F, G)-*integral* of W over Γ is defined by the equation:

$$\int_\Gamma W d_{(F,G)} z = F(z_1)\mathrm{Re}\int_\Gamma \frac{\bar{G}W\,dz}{F\bar{G} - \bar{F}G} - G(z_1)\mathrm{Re}\int \frac{\bar{F}W\,dz}{F\bar{G} - \bar{F}G} \quad (3)$$

W is called (F, G)-*integrable* in a domain $\mathbf{D}$ if its (F, G)-integral over a closed curve Γ vanishes whenever Γ is homologous to zero in $\mathbf{D}$.

THEOREM. *A continuous function is (F, G)-integrable if and only if it is an (F, G)-derivative;*

$$\int_{z_0}^{z_1} \dot{w} d_{(F,G)} z = w(z_1) - \varphi(z_0)F(z_1) - \psi(z_0)G(z_1). \quad (4)$$

A generating pair (F_1, G_1) is called a *successor* of (F, G) if (F, G)-derivatives are (F_1, G_1)-pseudo-analytic, and (F_1, G_1)-pseudo-analytic functions are (F, G)-integrable. (F, G) is then called a *predecessor* of (F_1, G_1).

THEOREM. *Let (F, G) be a generating pair in $\mathbf{D}_0$. In every bounded domain $\mathbf{D}_1$, $\bar{\mathbf{D}}_1 \subset \mathbf{D}_0$, (F, G) possesses a successor and a predecessor.*

A sequence of generating pairs $\{(F_\nu, G_\nu)\}$, $\nu = 0, \pm 1, \pm 2, \ldots$, is called a *generating sequence* if $(F_{\nu+1}, G_{\nu+1})$ is a successor of (F_ν, G_ν). A generating pair (F, G) in $\mathbf{D}_0$ can be embedded (in every bounded domain $\mathbf{D}_1$, $\bar{\mathbf{D}}_1 \subset \mathbf{D}_0$) in a generating sequence $\{(F_\nu, G_\nu)\}$ such that $(F_0, G_0) = (F, G)$. With respect to such a sequence an (F, G)-pseudo-analytic function $w(z)$ has

derivatives of all orders:

$$w^{(0)} = w, \quad w^{(n+1)} = d_{(F_n, G_n)}\, w^{(n)}/dz, \quad n = 1, 2, \ldots \qquad (5)$$

The sequence $\{w^{(n)}(z_0)\}$ determines the function $w(z)$ uniquely.

Two generating pairs, (F, G) and $(\tilde{F}, \tilde{G})$ are called *equivalent* if $\tilde{F} = a_{11}F + a_{12}G$, $\tilde{G} = a_{21}F + a_{22}G$, the a_{jk} being real constants. A generating sequence $\{(F_\nu, G_\nu)\}$ is called periodic if there exists an integer $\mu > 0$ such that $(F_{\nu+\mu}, G_{\nu+\mu})$ is equivalent to (F_ν, G_ν). In a bounded domain every generating pair can be embedded in a non-periodic generating sequence. I conjecture, but have not proved, that generating pairs cannot be embedded in periodic sequences, except in special cases.

Remark: Since (F,G)-differentiability is a conformally invariant property, pseudo-analyticity can be defined on a Riemann surface.

3. *Formal powers.*—A generating pair (F, G) in $\mathbf{E}_0$ is called a generating pair in $\mathbf{E}$ if $F(z)$ and $G(z)$ are of class $\mathfrak{C}^1$ at $z = \infty$, and $\operatorname{Im}\{\overline{F(\infty)}G(\infty)\} > 0$.

THEOREM. *A generating pair (F, G) in $\mathbf{E}$ can be embedded in a generating sequence $\{(F_\nu, G_\nu)\}$ in $\mathbf{E}_0$ in such a way that the functions $F_\nu(z)$, $G_\nu(z)$ are continuous at $z = \infty$. The generating pairs (F_ν, G_ν) are determined uniquely, up to equivalences.*

In this and the next section we consider a fixed generating pair (F, G) in $\mathbf{E}$ embedded in a sequence described in the theorem. This involves no serious loss of generality, since given a generating pair in $\mathbf{D}_0 \subset \mathbf{E}_0$, we can find a generating pair in $\mathbf{E}$ which coincides with the given one on a preassigned closed subdomain $\bar{\mathbf{D}}_1 \subset \mathbf{D}_0$.

Let a and z_0 be complex numbers, $a \neq 0$, and r a real rational number. If $r \neq 0$, set $|r| = p/q$, p and q being relatively prime positive integers; if $r = 0$, set $p = 0$, $q = 1$. An (F_ν, G_ν)-pseudo-analytic function $w(z)$, q-valued and regular for $0 < |z - z_0| < +\infty$ and such that $w(z) \sim a(z - z_0)^r$, $z \to z_0$, and $w(z) = 0(|z|^r)$, $z \to \infty$, is called a (*global*) *formal power* and is denoted by $Z_\nu^{(r)}(a, z_0; z)$. The corresponding function of the second kind is denoted by ${}_*Z_\nu^{(r)}(a, z_0; z)$. For $a = 0$, we set $Z_\nu^{(r)}(0, z_0; z) = 0$.

THEOREM. *The formal powers $Z_\nu^{(r)}(a, z_0; z)$ exist and are unique (for all ν, r, a, z_0). If $a \neq 0$, then the mapping $\omega(z) = {}_*Z_\nu^{(r)}(a, z_0, z)$ is a homeomorphism of the q-times covered z-plane (with branch-points at z_0 and ∞) onto the p-times covered ω-plane (with branch-points at 0 and ∞). $Z_\nu^{(r)}(a, z_0, z)$ is a continuous function of z_0. The formal powers satisfy the relations:*

$$Z_\nu^{(r)}(\lambda a + \mu b, z_0; z) = \lambda Z_\nu^{(r)}(a, z_0; z) + \mu Z_\nu^{(r)}(a, z_0; z) \qquad (6)$$

where λ and μ are real constants, and

$${}_*Z_\nu^{(0)}(a, z_0; z) = a, \; d_{(F_\nu, G_\nu)}\, Z_\nu^{(r)}(a, z_0; z)/dz = rZ_{(\nu+1)}^{(r-1)}(a, z_0; z). \qquad (7)$$

It follows from (7) that once the sequence $\{(F_\nu, G_\nu)\}$ is known, the formal powers $Z_\nu^{(n)}$, $n = 1, 2, \ldots$, can be obtained by successive integrations.

The same procedure can be carried out starting with an arbitrary generating sequence defined in a bounded domain $\mathbf{D}_0$. In this way we obtain *local formal powers*, which have some properties in common with the global powers. (Local formal powers with negative and/or fractional exponents can also be defined, but not uniquely.)

4. *Applications of Formal Powers.*—With the aid of the global formal powers we can reestablish all results from I §§5-13 in the present more general circumstances. Thus we obtain Taylor and Laurent expansions for single-valued and multiple-valued functions, interpolation theorems for rational (regular but for a finite number of poles) pseudo-analytic functions, the analogs of the fundamental theorem of algebra, of Cauchy's integral formula,[4] of Weierstrass' convergence theorem, of Runge's approximation theorem, etc. We also state a result which describes the totality of single-valued pseudo-analytic functions.

Let $w(z)$ be single-valued and (F, G)-pseudo-analytic in a domain $\mathbf{D}$, and let $f(z)$ be single-valued and analytic in this domain. We say that $w(z)$ and $f(z)$ are *similar* if there exists a continuous function $\chi(z)$ in $\bar{\mathbf{D}}$, such that $w(z) = e^{\chi(z)}f(z)$.

THEOREM. *Every single-valued analytic function $f(z)$, $z \in \mathbf{D}$, has a similar (F, G)-pseudo-analytic function $w(z)$, and vice versa. If one of the functions is given, the other may be chosen so that $[w(z)/f(z)] \to 1$ as $z \to z_0$, z_0 being a prescribed point of $\bar{\mathbf{D}}$.*

This theorem implies, among other things, the existence of pseudo-analytic functions with infinitely many prescribed isolated zeros and poles.

5. *Pseudo-analytic Functions and Partial Differential Equations.*—In order to express pseudo-analyticity by differential equations we associate with every generating pair (F, G) the functions

$$a = -(\bar{F}G_z - F_z\bar{G})/(F\bar{G} - \bar{F}G), \quad b = (FG_z - F_zG)/(F\bar{G} - \bar{F}G), \tag{8}$$

$$A = -(\bar{F}G_z - F_z\bar{G})/(F\bar{G} - \bar{F}G), \quad B = (FG_z - F_zG)/(F\bar{G} - \bar{F}G), \tag{9}$$

$$\alpha - i\beta = 4\,(\bar{F}/F)(FG_z - F_zG)/(F\bar{G} - \bar{F}G), \tag{10}$$

α and β being real-valued.

LEMMA. *Let $\mathbf{D}_0$ and $\mathbf{D}_1$ be domains such that $\bar{\mathbf{D}}_1 \subset \mathbf{D}_0 \subset \mathbf{E}_0$.* (a) *Given two functions of class $\mathfrak{C}$, $a(z)$ and $b(z)$, defined in $\mathbf{D}_0$, there exists a generating pair (F, G) in $\mathbf{D}_1$ such that* (8) *holds.* (b) *Given two real-valued functions of class $\mathfrak{C}$, $\alpha(z)$ and $\beta(z)$, defined in $\mathbf{D}_0$, there exists a generating pair (F, G) in $\mathbf{D}_1$ such that* (10) *holds.*

The statements (a) and (b) remain valid for $\mathbf{D} = \mathbf{D}_0 = \mathbf{E}_0$ if $a, b, \alpha, \beta = O(|z|^{-2})$, $z \to \infty$.

THEOREM. *A function $w(z) = \varphi(z)F(z) + \psi(z)G(z)$ of class $\mathfrak{C}^1$ is (F, G)-pseudo-analytic if and only if*

$$w_{\bar{z}} = aw + b\bar{w} \tag{11}$$

i.e., if and only if

$$F\varphi_z + G\psi = 0. \tag{12}$$

If this condition is satisfied, then

$$\dot{w} = w_z - Aw - B\bar{w} \equiv F\varphi_z + G\psi_z \tag{13}$$

and

$$(\dot{w})_{\bar{z}} = a\dot{w} - B\bar{w}. \tag{14}$$

(Note that (14) and part (a) of the lemma imply the existence of successors.)

Eliminating from (12) the function φ we obtain the equation

$$\psi_{xx} + \psi_{yy} + \alpha\psi_x + \beta\psi_y = 0 \tag{15}$$

where α and β are given by (10).

Example: Let $\sigma(z)$ be a positive function of class $\mathfrak{C}^1$, and set $(F, G) = (\sigma^{-1/2}, i\sigma^{1/2})$. Then the (F, G)-pseudo-analytic functions of the first (second) kind are the reduced pseudo-analytic (pseudo-analytic) functions of I. If $\sigma = \sigma_1(x)\sigma_2(y)$, then the functions of the second kind are sigma-monogenic.[2] In this case a generating sequence can be obtained by setting $(F_{2\nu}, G_{2\nu}) = (F, G)$, $(F_{2\nu+1}, G_{2\nu+1}) = (\tilde{\sigma}^{-1/2}, i\tilde{\sigma}^{1/2})$, where $\tilde{\sigma} = \sigma_2(y)/\sigma_1(x)$.

Consider now the general linear homogeneous partial differential equation (with real coefficients):

$$A_{11}\psi_{xx} + 2A_{12}\psi_{xy} + A_{22}\psi_{yy} + B_1\psi_x + B_2\psi_y + \Gamma\psi = 0. \tag{16}$$

We assume that the coefficients A_{jk} are of class $\mathfrak{C}^1$, the coefficients B_j, of class $\mathfrak{C}$, and that the ellipticity condition $(A_{11}A_{22} - A_{12}^2 > 0)$ is satisfied in the domain considered.[5] By introducing new independent variables (16) can be reduced to the form

$$\psi_{xx} + \psi_{yy} + \alpha\psi_x + \beta\psi_y + \gamma\psi = 0. \tag{17}$$

If $\gamma \equiv 0$, this equation is at the form (15) and it follows from statement (b) of the lemma that, for an appropriate choice of (F, G), the solutions of our equation will be the imaginary parts of (F, G)-pseudo-analytic functions of the second kind. If $\gamma \not\equiv 0$, we consider equation (17) in a domain where it possesses a positive solution ψ_0, and introducing the new unknown function $\tilde{\psi} = \psi/\psi_0$ reduce the equation to the form (15).

Consider next the general system of linear homogeneous partial differential equations (with real coefficients).

$$A_{j1}\varphi_x + A_{j2}\varphi_y + B_{j1}\psi_x + B_{j2}\psi_y + \Gamma_{j1}\varphi + \Gamma_{j2}\psi = 0, j = 1, 2. \tag{18}$$

This system is elliptic if the roots of the equation $\det(A_{jk} - \lambda B_{jk}) = 0$

are complex. We assume that this condition is satisfied in the domain considered, and that the coefficients A_{jk}, B_{jk} are of class $\mathfrak{C}^1$, the coefficients Γ_{jk} of class $\mathfrak{C}$. By introducing new independent variables (18) can be reduced to the form

$$\varphi_x = \tau\psi_x + \sigma\psi_y + \gamma_{11}\varphi + \gamma_{12}\psi, \quad \varphi_y = -\sigma\psi_x + \tau\psi_y + \gamma_{21}\varphi + \gamma_{22}\psi \tag{19}$$

where σ has constant sign. Without loss of generality we assume that $\sigma > 0$. Set $\gamma^2 = \gamma_{11}^2 + \gamma_{12}^2 + \gamma_{21}^2 + \gamma_{22}^2$. If $\gamma \equiv 0$, equations (19) are equivalent to equation (12) for $F = 1$, $G = -\tau + i\sigma$. Solutions (φ, ψ) are then real and imaginary parts of (F, G)-pseudo-analytic functions of the second kind. If $\gamma \not\equiv 0$, we reduce (19) to a system of the same form with $\gamma \equiv 0$, by introducing new unknown functions $\tilde{\varphi}$, $\tilde{\psi}$ such that $\varphi = \varphi_1\tilde{\varphi} + \varphi_2\tilde{\psi}$, $\psi = \psi_1\varphi + \psi_2\psi$, where (φ_1, ψ_1), (φ_2, ψ_2) are fixed solutions of (19) with $\varphi_1\psi_2 - \varphi_2\psi_1 > 0$. The existence of such solutions follows easily from statement (a) of the lemma.

On the other hand, we can introduce new unknown functions $u = \varphi - \tau\psi$, $v = \sigma\psi$. Then $w = u + iv$ satisfies an equation of the form (11). It follows from (a) that, for an appropriate choice of (F, G), solutions (u, v) are real and imaginary parts of (F, G)-pseudo-analytic functions of the first kind.

Thus the theory of pseudo-analytic functions bears the same relationship to the general theory of elliptic equations as the theory of analytic functions does to that of the Laplace equation.

[1] Bers, L., *Proc. Natl. Acad. Sci.*, **36**, 130–136 (1950).

[2] Bers, L., and Gelbart, A., *Quart. Appl. Math.*, **1**, 168–188 (1943); *Trans. Am. Math. Soc.*, **56**, 67–93 (1944); *Ann. Math.*, **48**, 342–357 (1947); Bers, L., *Am. J. Math.*, **72**, 705–712 (1950). For a generalization see Lukomskaya, M. A., *Doklady Akad. Nauk*, **73**, 895–888 (1950).

[3] Reported by I. G. Petrovskii, *Uspekhi Mat. Nauk*, **1**, 44–70 (1946).

[4] For other generalizations of Cauchy's formula see Bergman, S., *Trans. Am. Math. Soc.*, **62**, 452–497 (1947); Položii, G. N., *Mat. Sbornik*, **24** (66), 375–384 (1949); Šabat, B. V., *Doklady Akad. Nauk*, **69**, 305–308 (1949).

[5] We do not consider here the interesting case of equations possessing parabolic lines; cf. reference 2, and Weinstein, A., *Trans. Am. Math. Soc.*, **63**, 342–354 (1948); *Quart. Appl. Math.*, **5**, 429–444 (1948).

THE HARMONIC OPERATOR FOR EXTERIOR DIFFERENTIAL FORMS*, †

BY MATTHEW P. GAFFNEY

NORTHWESTERN UNIVERSITY

Communicated by M. H. Stone, October 30, 1950

Harmonic integrals were introduced by Hodge.[1] Kodaira[2, 3] and Bidal and de Rham[4, 5] independently introduced the harmonic operator $\Delta(= d\delta + \delta d)$; one of their principal uses for Δ was to study harmonic integrals. Δ is defined for exterior differential forms on a Riemannian manifold of class C^k; for functions it reduces to the Beltrami operator. If we properly select its domain, Δ can be regarded as a linear transformation on a Hilbert space—the Hilbert space is obtained by suitably completing the class of continuous square integrable forms—and the first object of our work is to develop the properties of Δ from this point of view (thus avoiding the use of integral equation theory). Our second—and principal—object is to extend our knowledge of Δ in the non-compact case (in which the integral equation theory does not work well). Here the most important question is whether or not the closure, $\bar{\Delta}$, is self-adjoint; for only when it is can the powerful spectral theorem be applied.[6]

The domain of d (respectively δ) we take to be the set of all α in C^1 such that both α and $d\alpha$ (respectively $\delta\alpha$) are square integrable. Similarly, the domain of Δ is the set of those α such that α, $d\alpha$, and $\delta\alpha$ are all in the domains of d and δ. With this definition Δ will not, in general, be symmetric; we are therefore led to restricting the class of manifolds considered. We demand that d and δ (with the domains provided) be adjoint $[(d\alpha, \beta) = (\alpha, \delta\beta)]$. When this condition is satisfied we say that the manifold has "negligible boundary." Δ is then easily seen to be symmetric; our main result is that for manifolds with negligible boundary $\bar{\Delta}$ is self-adjoint.

If a manifold is compact it has negligible boundary. Another sufficient condition is that there exist a function s with the following properties (which are suggested by a consideration of the function equal to the distance from a variable point to the origin in euclidean space):

I. Outside some fixed compact set K the function s is of class C^1.

II. There exists a number t_0 such that for all $t \geqq t_0$ the set $S_t = [P: s(P) \leqq t]$ is a compact set containing K.

III. Outside K the gradient of s is never zero and the magnitude of the gradient is bounded.

These properties can be weakened. We can also give a different sufficient condition which applies, for example, to the euclidean plane with a point removed.

The first step in proving self-adjointness is to prove that $\bar{d} = \delta^*$: for

every α in the domain of δ^* we must find a sequence $<\alpha_m>$ of forms in the domain of d such that in the L_2 norm [norm $\alpha = \sqrt{(\alpha, \alpha)}$] α_m approaches α and $d\alpha_m$ approaches $\delta^*\alpha$. We first localize the problem to a single coordinate neighborhood by the use of a partition of unity, $1 = \Sigma_i \phi_i$. To use the partition we must prove that if α is in the domain of δ^* so also is $\phi_i\alpha$. (While this follows readily, the corresponding fact for Δ^* offers considerable difficulty.) To produce the approximating sequence $<\alpha_m>$ we use the Friedrichs mollifier;[7, 8] this is an integral operator on E^n whose kernel is an approach to the Dirac delta function. On the real line (for example) the kernel is defined by $j_\epsilon(x - y) = \frac{1}{\epsilon} f\left(\frac{x - y}{\epsilon}\right)$, where $f(t)$ is a C^∞ non-negative function which is zero outside the interval $[-1, 1]$ and whose integral over this interval is 1. For a form α (which by localization vanishes outside a coordinate neighborhood) we first fix the coordinate system and then apply the mollifier to the individual coefficients. Denoting this operator by J_ϵ we prove that if α is an L_2 form then $J_\epsilon\alpha$ approaches α. Furthermore, if α is also in the domain of δ^*, then $dJ_\epsilon\alpha = J_\epsilon\delta^*\alpha$ and consequently approaches $\delta^*\alpha$; this proves that $\bar{d} = \delta^*$, and it follows that $\bar{\delta} = d^*$. By a standard theorem $d\bar{\delta}$ and $\bar{\delta}d$ are therefore self-adjoint, and with the aid of the relations $dd = \delta\delta = 0$ it can be shown that $d\bar{\delta} + \bar{\delta}\bar{d}$ is self-adjoint.

We must next show that $\bar{\Delta} = d\bar{\delta} + \bar{\delta}\bar{d}$. First we prove that if α is in the domain of d^* then $\delta J_\epsilon\alpha$ approaches $d^*\alpha$; since δ and J_ϵ do not commute this requires a special argument, but the methods of Friedrichs provide the necessary tools. Then if α is in the domain of δd, we have $\delta d J_\epsilon\alpha = \delta J_\epsilon \bar{d}\alpha$ approaches $\delta d\alpha$—note the important role played by the commutativity of d and J_ϵ. Next it is shown that if β is in the domain of $d\bar{\delta} + \bar{\delta}\bar{d}$ so also is $\phi_i\beta$; this proof uses the facts concerning δJ_ϵ and also the inequality we give below. Finally, techniques are developed to prove that $\Delta J_\epsilon(\phi_i\beta)$ approaches $(d\bar{\delta} + \bar{\delta}\bar{d})(\phi_i\beta)$.

The inequality mentioned is of independent interest. For a particular coordinate neighborhood denote by $D(\alpha)$ the integral of the sum of the squares of all possible first partial derivatives of the coefficients of α. Then for every point P there is a coordinate neighborhood K such that for all C^2 forms vanishing outside K

$$D(\alpha) \leqq C[(d\alpha, d\alpha) + (\delta\alpha, \delta\alpha) + (\alpha, \alpha)],$$

where C is a constant depending on K and the coordinate system.

When the manifold is sufficiently differentiable every form α such that $\bar{\Delta}\alpha = 0$ is shown to be of class C^2 and therefore harmonic. The method of proof is to represent α in the range of an integral operator whose kernel is constructed (see reference 4) with the aid of the parametrix $1/r^{n-2}$. By iteration the kernel can be made square integrable.

For compact manifolds we use a method of Rellich's[9] to prove directly, without recourse to a representation as an integral operator, that the inverse of Δ is completely continuous.

Consider the three linear subspaces of p-forms consisting of, respectively, the harmonic forms, the range of d and the range of δ. The self-adjointness of $\bar{d}\delta + \bar{\delta}d$ easily implies that the closures of these three spaces span the Hilbert space of p-forms. Together with the other results this leads to the standard decomposition theorem (method of orthogonal projections) for compact manifolds and a variation thereof for non-compact manifolds. It is interesting that a simple theorem of Hilbert space theory (that for a self-adjoint operator the closures of the range and the set mapped into zero span the Hilbert space) appears as a tool for proving global theorems.

* Paper done under contract with the Office of Naval Research, N6ori 20, T.O. 25, NR 043-083; Mathematics Department, University of Chicago. To be submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Chicago.

† The subject here presented was suggested to me in outline by Prof. M. H. Stone. In addition, I received valuable aid from Prof. K. O. Friedrichs.

[1] Hodge, W. V. D., "A Dirichlet Problem for Harmonic Functionals, with Applications to Analytic Varieties," *Proc. London Math. Soc.*, **36,** ser. 2, 72–95 (1932).

[2] Kodaira, K., "Über die harmonischen Tensorfelder in Riemannschen Mannigfaltigkeiten," *Proc. Imp. Acad. Tokyo*, **20,** 186–198, 257–261, 353–358 (1944).

[3] Kodaira, K., "Harmonic Fields in Riemannian Manifolds (Generalized Potential Theory)," *Ann. Math.*, **50,** 587–665 (1949).

[4] Bidal, P., and de Rham, G., "Les formes différentielles harmoniques," *Comment. Math. Helv.*, **19,** 1–49 (1946).

[5] de Rham, G., and Kodaira, K., "Harmonic Integrals," mimeographed notes, Institute for Advanced Study (1950).

[6] We use "self-adjoint" in the Hilbert space sense, which in the real case means that the operator is not only symmetric but that no extension of it is symmetric.

[7] Friedrichs, K., "On Differential Operators in Hilbert Spaces," *Am. J. Math.*, **61,** 523–544 (1939).

[8] Friedrichs, K., "The Identity of Weak and Strong Extensions of Differential Operators, *Trans. Am. Math. Soc.*, **55,** 132–151 (1944).

[9] Rellich, F., *Ein Satz über mittlere Konvergenz*, Göttingen Nachrichten, 1930, pp. 30–35.

ON THE ANALYSIS OF REPRESENTATIONS OF THE LINEAR GROUP

BY FRANCIS D. MURNAGHAN

INSTITUTO TECNOLÓGICO DE AERONÁUTICA, RIO DE JANEIRO

Communicated October 16, 1950

There is associated with each partition of any positive integer m into not more than n parts an irreducible, homogeneous, rational integral representation of degree m of the full linear group of dimension n and all such representations of the full linear group are furnished by this association. If a is a typical matrix of the linear group of dimension n we denote by $A = a(\lambda)$ the matrix which represents a in the irreducible representation which is associated with the partition $(\lambda) = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ of m, $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k > 0$, $k \leqslant n$, and we denote by $\{\lambda\}$ the trace of A. $\{\lambda\}$ is a polynomial function of the power sums $s_1, s_2, \ldots, s_m$ of the characteristic numbers of a which is known (in honor of I. Schur) as an S-function and which is said to be of weight m. If (μ) is any partition of any positive integer p into not more than N parts (where N is the dimension of the various matrices A) the irreducible representation of the full linear group of dimension N which is associated with (μ) induces a representation, in general reducible, of the group of matrices A and we denote the characters of this representation by the symbol $\{\lambda\} \otimes \{\mu\}$. Thus $\{\lambda\} \otimes \{\mu\}$ is the S-function of the power sums $S_1, S_2, \ldots, S_p$ of the characteristic numbers of the matrices $A = a(\lambda)$ which is associated with the partition (μ) of p; it is, accordingly, a polynomial function of weight mp of $s_1, s_2, \ldots, s_m$ and, hence, a linear combination, with non-negative integral coefficients, of S-functions of weight mp. The "new product" (so called by Littlewood[1]) $\{\lambda\} \otimes \{\mu\}$ of S-functions is of central importance in the theory of algebraic invariants. Littlewood[1] has given some three methods for carrying out this new multiplication of S-functions but he confesses that, even in the case where the partition (μ) contains only one element, "many difficulties arise and it has not so far been possible to overcome them." In the present note we announce some results obtained in studying this problem.

In the first place we remark that $\{\lambda\} \otimes \{\mu\}$ may be read off from the character table of the symmetric group on mp letters. For example, $\{31^2\} \otimes \{2\} = \frac{1}{2}(S_1^2 + S_2)$ where $S_1 = \frac{1}{20}(s_1^5 - 5s_1s_2^2 + 4s_5)$, $S_2 = \frac{1}{20}(s_2^5 - 5s_2s_4^2 + 4s_{10})$. S_1^2 is the square of $\{31^2\}$ which is available in the literature[2] and S_2 is furnished by three columns of the character table of the symmetric group on 10 letters (note that we only require a small portion of the character table and that this is the portion which is relatively easy to

calculate). We obtain $S_1^2 = \{62^2\} + \{621^2\} + \{61^4\} + \{532\} + \{531^2\} + 2\{52^21\} + 2\{521^3\} + \{51^5\} + \{4^22\} + \{4^21^2\} + 2\{4321\} + 2\{431^3\} + \{42^3\} + \{42^21^2\} + \{421^4\} + \{3^22^2\} + \{3^221^2\} + \{3^21^4\}$, $S_2 = \{62^2\} - \{621^2\} + \{61^4\} - \{532\} + \{531^2\} - \{51^5\} + \{4^22\} - \{4^21^2\} + \{42^3\} - \{42^21^2\} + \{421^4\} - \{3^22^2\} + \{3^221^2\} - \{3^21^4\}$ so that $\{31^2\} \otimes \{2\} = \{62^2\} + \{61^4\} + \{531^2\} + \{52^21\} + \{521^3\} + \{4^22\} + \{4321\} + \{431^3\} + \{42^3\} + \{421^4\} + \{3^221^2\}$. At the same time we obtain $\{31^2\} \otimes \{1^2\} = \frac{1}{2}(S_1^2 - S_2) = \{621^2\} + \{532\} + \{52^21\} + \{521^3\} + \{51^5\} + \{4^21^2\} + \{4321\} + \{431^3\} + \{42^21^2\} + \{3^22^2\} + \{3^21^4\}$ which is, as it should be, the sum of the conjugates of the terms in $\{31^2\} \otimes \{2\}$. This method is quite practicable for all partitions (λ) of m and (μ) of p where $mp \leqslant 12$. For higher values of mp it becomes impracticable since the character table of the symmetric group on r letters is an unwieldy object if $r > 12$ (it being a 231×231 table if $r = 16$ and a 627×627 table if $r = 20$).

We denote by $\{m\}$ the partition of m which contains only one element $\lambda_1 = m$. Then the formula $\{m\} \otimes \{2\} = \{2m\} + \{2m-2, 2\} + \ldots + \{m^2\}$, if m is even, and $\{m\} \otimes \{2\} = \{2m\} + \{2m-2, 2\} + \ldots + \{m+1, m-1\}$, if m is odd, has been furnished by Littlewood.[1] Since $\{m\}\{m\} = \{2m\} + \{2m-1, 1\} + \ldots + \{m, m\}$ this implies that $\{m\} \otimes \{1^2\} = \{2m-1, 1\} + \{2m-3, 3\} + \ldots + \{m+1, m-1\}$, if m is even, and $\{m\} \otimes \{1^2\} = \{2m-1, 1\} + \{2m-3, 3\} + \ldots + \{m^2\}$, if m is odd. To these we add the following two results which are valid for any positive integer $m > 2$:

I(a). $$\{m-1, 1\} \otimes \{2\} = \sum_j \{2m-2j, 2j\} + \sum_j \{2m-j-2, j+1, 1\} + \sum_j \{2m-2j-2, 2j, 2\} + \sum_j \{2m-2j-1, 2j-1, 1^2\}.$$

Each of the four summations starts out with $j = 1$ and continues as long as the partitions are in natural (non-ascending) order. Thus $\{51\} \otimes \{2\} = \{10.2\} + \{921\} + \{91^3\} + \{84\} + \{831\} + \{82^2\} + \{741\} + \{731^2\} + \{6^2\} + \{651\} + \{642\} + \{5^21^2\}$.

I(b). $$\{m-1, 1\} \otimes \{1^2\} = \sum_j \{2m-2j-1, 2j+1\} + \sum_j \{2m-j-1, j, 1\} + \sum_j \{2m-2j-3, 2j+1, 2\} + \sum_j \{2m-2j-2, 2j, 1^2\}.$$

Thus $\{61\} \otimes \{1^2\} = \{10.1^2\} + \{93\} + \{921\} + \{831\} + \{821^2\} + \{75\} + \{741\} + \{732\} + \{651\} + \{641^2\} + \{5^22\}$.

On adding the formulas I(a) and I(b) we obtain

I(*c*). $\{m-1,1\}^2 = \{2m-2,1^2\} + \sum_j \{2m-j-1,j+1\} + 2\sum_j \{2m-j-2,j+1,1\} + \sum_j \{2m-j-3,j+1,2\} + \sum_j \{2m-j-2,j,1^2\}.$

Thus $\{61\}^2 = \{12.2\} + \{12.1^2\} + \{11.3\} + 2\{11.21\} + \{11.1^3\} + \{10.4\} + 2\{10.31\} + \{10.2^2\} + \{95\} + 2\{941\} + \{932\} + \{931^2\} + \{86\} + 2\{851\} + \{842\} + \{841^2\} + \{7^2\} + 2\{761\} + \{752\} + \{751^2\} + \{6^22\} + \{6^21^2\}.$

II(*a*). $\{m-2,2\} \otimes \{2\} = \{2m-4,2^2\} + \sum_j \{2m-2j-2,2j+2\} + \sum_j \{2m-j-4,j+3,1\} + \sum_j \{2m-2j-3,2j+1,2\} + 2\sum_j \{2m-2j-4,2j+2,2\} + \sum_j \{2m-2j-3,2j+1,1^2\} + \sum_j \{2m-j-6,j+3,3\} + \sum_j \{2m-j-5,j+2,2,1\} + \sum_j \{2m-2j-6,2j+2,4\} + \sum_j \{2m-2j-5,2j+1,3,1\} + \sum_j \{m-2j-4,2j,2^2\} - \{(m-1)\,2\}.$

II(*b*). $\{m-2,2\} \otimes \{1^2\} = \{2m-4,3,1\} + \{2m-5,3,2\} - \{(m-1)^22\} + \sum_j \{2m-2j-3,2j+3\} + \sum_j \{2m-j-4,j+3,1\} + 2\sum_j \{2m-2j-5,2j+3,2\} + \sum_j \{2m-2j-4,2j+2,2\} + \sum_j \{2m-2j-4,2j+2,1^2\} + \sum_j \{2m-j-5,j+2,3\} + \sum_j \{2m-j-4,j+1,2,1\} + \sum_j \{2m-2j-7,2j+3,4\} + \sum_j \{2m-2j-6,2j+2,3,1\} + \sum_j \{2m-2j-5,2j+1,2^2\}.$

II(*c*). $\{m-2,2\}^2 = \{2m-4,3,1\} + \{2m-4,2^2\} + 2\{2m-5,3,2\} + \{2m-5,2^2,1\} + \{2m-6,3^2\} - 2\{(m-1)^22\} + \sum_j \{2m-j-3,j+3\} + 2\sum_j \{2m-j-4,j+3,1\} + 3\sum_j \{2m-j-5,j+3,2\} + \sum_j \{2m-j-4,j+2,1^2\} + 2\sum_j \{2m-j-6,j+3,3\} + 2\sum_j \{2m-j-5,j+2,2,1\} + \sum_j \{2m-j-7,j+3,4\} + \sum_j \{2m-j-6,j+2,3,1\} + \sum_j \{2m-j-5,j+1,2^2\}.$

If the integer m of which (λ) is a partition is even and if we denote by (λ^*) the partition of m which is conjugate to, or associated with, (λ) then $\{\lambda^*\} \otimes \{\mu\}$ is obtained by starring all the terms which appear in the analysis of $\{\lambda\} \otimes \{\mu\}$ while, if m is odd, $\{\lambda^*\} \otimes \{\mu\}$ is obtained by starring all

the terms which appear in the analysis of $\{\lambda\} \otimes \{\mu^*\}$. Thus the formulas we have given serve to determine $\{\lambda\} \otimes \{\mu\}$ when (μ) is (2) or (1^2) and (λ) is any partition of 3, 4 or 5 with the single exception of the self-associated partition (31^2) of 5. Since we have given above $\{31^2\} \otimes \{2\}$ and $\{31^2\} \otimes \{1^2\}$ we are in possession of $\{\lambda\} \otimes \{2\}$ and $\{\lambda\} \otimes \{1^2\}$ when (λ) is any partition of 3, 4 or 5 and, generally, if $(\lambda) = (m-1, 1), (m-2, 2), (21^{m-2}), (2^2 1^{m-4})$.

The method employed by Littlewood and his coworkers[1, 3] depends on the following fact. If $\{\lambda\}$ is any S-function the derivative of $\{\lambda\}$ with respect to s_1 is the sum of the S-functions obtained by subtracting 1, in turn, from each element of (λ). We denote this derivative of $\{\lambda\}$ with respect to s_1 by $\{\lambda\}_1$ and it follows at once that $[\{\lambda\} \otimes \{\mu\}]_1 = [\{\lambda\} \otimes \{\mu\}_1] + [\{\lambda\}_1 \otimes \{\mu\}]$. From a knowledge of the derivative of $\{\lambda\} \otimes \{\mu\}$ with respect to s_1 Littlewood tries to obtain $\{\lambda\} \otimes \{\mu\}$ but since $\{\lambda\} \otimes \{\mu\}$ involves, in addition to s_1, the variables $s_2, s_3, \ldots$, this passage cannot be made with certainty. In Littlewood's words "at times a choice of certain alternatives is presented. Choosing the correct partitions the work develops rapidly to the correct conclusion. But if a wrong partition is chosen the working goes further and further astray until the error is apparent . . . much time can be lost in following false tracks." These false tracks may be avoided when we observe that the relation $[\{\lambda\} \otimes \{\mu\}]_1 = [\{\lambda\} \otimes \{\mu\}_1] + [\{\lambda\}_1 \otimes \{\mu\}]$ is only the first of a series of such relations. If q is any positive integer the derivative of $\{\lambda\}$ with respect to s_q/q is the sum of the S-functions obtained by subtracting q in turn from each element of (λ) (it being understood that (λ) is extended by the addition of an indefinite number of zeros and that disordered partitions are re-ordered in the usual way). We denote this derivative of $\{\lambda\}$ with respect to s_q/q by the symbol $\{\lambda\}_q$ and we have the following formulas for $[\{\lambda\} \otimes \{\mu\}]_q, q = 2, 3, \ldots$

$$[\{\lambda\} \otimes \{\mu\}]_2 = [\{\lambda\} \otimes \{\mu\}_1] + [\{\lambda\} \otimes \{\mu\}_2] + \{\{\lambda\}_1 \otimes [\{2\} - \{1^2\}]\};$$

$$[\{\lambda\} \otimes \{\mu\}]_3 = [\{\lambda\} \otimes \{\mu\}_1] + [\{\lambda\} \otimes \{\mu\}_3] + \{\{\lambda\}_1 \otimes [\{3\} - \{21\} + \{1^3\}]\};$$

$$[\{\lambda\} \otimes \{\mu\}]_4 = [\{\lambda\} \otimes \{\mu\}_1] + [\{\lambda\} \otimes \{\mu\}_4] + \{[\{\lambda\} \otimes \{\mu\}_2] + \{\lambda\}_2 \otimes [\{2\} - \{1^2\}]\} + \{\{\lambda\}_1 \otimes [\{4\} - \{31\} + \{21^2\} - \{1^4\}]\}$$

and so on. For example,

$$[\{2\} \otimes \{2^2\}]_1 = [\{2\} \otimes \{21\}] + \{1\}; \quad [\{2\} \otimes \{2^2\}]_2 = \{2\} \otimes \{21\} + \{\{2\} \otimes [\{2\} - \{1^2\}]\} + [\{2\} - \{1^2\}];$$

$$[\{2\} \otimes \{2^2\}]_3 = -\{2\}[\{3\} - \{21\} + \{1^3\}];$$

$$[\{2\} \otimes \{2^2\}]_4 = \{2\} \otimes [\{2\} - \{1^2\}];$$

$$[\{2\} \otimes \{2^2\}]_5 = 0; \quad [\{2\} \otimes \{2^2\}]_6 = -\{2\};$$

$$[\{2\} \otimes \{2^2\}]_q = 0, \; q > 6.$$

In conclusion we remark that $\{2\} \otimes \{m\} = \sum \{2\lambda_1, \ldots, 2\lambda_k\}$, the summation being over all partitions (λ) of m; $\{2\} \otimes \{m-1, 1\}$ is obtained by

multiplying $\{2\} \otimes \{m-1\}$ by $\{2\}$ and subtracting $\{2\} \otimes \{m\}$. For the deeper partitions (μ) of m it seems simplest to use the method outlined above. Thus to evaluate $\{2\} \otimes \{m-2, 2\}$ we suppose that $\{2\} \otimes \{\mu\}$ has been evaluated for all partitions (μ) of any integer $< m$ and use the relations

$$[\{2\} \otimes \{m-2, 2\}]_1 = \{\{2\} \otimes [\{m-3, 2\} + \{m-2, 1\}]\}\{1\},$$
$$[\{2\} \otimes \{m-2, 2\}]_2 = \{2\} \otimes [\{m-3, 2\} + \{m-2, 1\}] + \{\{2\} \otimes [\{m-4, 2\} + \{m-2\}]\}\,[\{2\} - \{1^2\}]$$

and so on.[4]

[1] Littlewood, D. E., *Phil. Trans. Roy. Soc.*, Ser. A., **239**, 305–365 (1944).

[2] Murnaghan, F. D., *Am. J. Math.*, **60**, 44–65 (1938).

[3] Foulkes, H. O., *J. Lond. Math. Soc.*, **25**, 205–209 (1950).

[4] A paper by G. de B. Robinson, treating the problem discussed in the present note, has appeared in the current issue of the *Canadian Journal of Mathematics*.

THE CHARACTERS OF THE SYMMETRIC GROUP

BY FRANCIS D. MURNAGHAN

INSTITUTO TECNOLOGICO DE AERONÁUTICA, SÃO JOSÉ DOS CAMPOS, BRAZIL

Communicated November 25, 1950

If $(\lambda) = (\lambda_1, \lambda_2, \ldots, \lambda_k)$, $1 \leqslant k \leqslant m$, $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_k > 0$, is any partition of any positive integer m there is associated with (λ) an irreducible representation of the symmetric group S_m, i.e., the group of $m!$ permutations of m letters. The elements of S_m which have the same cycle structure, i.e., which contain α_1 unary cycles, α_2 binary cycles, ..., α_m cycles on m letters, belong to the same class $(\alpha) = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ and the character of any representation of S_m is a class function. The characters of the irreducible representations of S_m (each of which is associated with a partition of m) are known as the characters of S_m and Frobenius furnished formulas giving the character of $\{\lambda\} = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$, i.e., of the irreducible representation of S_m which is associated with the partition (λ) of m, as a function of $(\alpha) = (\alpha_1, \ldots, \alpha_m)$ for certain partitions of m. It seems to have escaped his attention, and that of subsequent writers on the subject, that these and other similar formulas can be procured in a purely mechanical manner involving little more than the trouble of writing them down; it is the object of this note to describe this mechanical process. An account of the theory underlying the process will appear in *Lectures on Matrices and Matrix Groups* to be published by the Centro de Pesquizas Fisicas, Rio de Janeiro.

We first define a process of differentiation of the symbols $\{\lambda\} = \{\lambda_1, \ldots, \lambda_k\}$. If p is any positive integer we denote by $\{\lambda_1, \ldots, \lambda_k\}_p$ the sum

$$\{\lambda_1, \ldots, \lambda_k\}_p = \{\lambda_1 - p, \lambda_2, \ldots, \lambda_k\} + \ldots, + \{\lambda_1, \lambda_2, \ldots, \lambda_{k-1}, \lambda_k - p\}$$

where any partition which is disordered by subtraction of p from one of its elements is supposed reordered in the usual way; we term $\{\lambda_1, \ldots, \lambda_k\}_p$ the derivative of $\{\lambda_1, \ldots, \lambda_k\}$ with respect to the integer p. Secondly, if we have any polynomial function of $\alpha_1, \ldots, \alpha_m$ we suppose it expressed in terms of the binomial coefficients $\alpha_2, \binom{\alpha_2}{2}, \binom{\alpha_2}{3}, \ldots, \alpha_3, \binom{\alpha_3}{2}, \ldots$, and we say that its integral with respect to α_2 is the function obtained by regarding $\alpha_1, \alpha_3, \ldots$ as constants and replacing any constant C by $C\alpha_2$, any constant times α_2 by that constant times $\binom{\alpha_2}{2}$, any constant times $\binom{\alpha_2}{2}$ by that constant times $\binom{\alpha_2}{3}$ and so on. Thus the integral of α_2 with respect to α_2, instead of being $\frac{1}{2}\alpha_2^2$, is $\frac{1}{2}\alpha_2(\alpha_2 - 1)$ and so on. Similarly the integral of any constant with respect to α_3 is that constant times α_3; the integral of any constant times α_3 with respect to α_3 is that constant times $\binom{\alpha_3}{2}$ and so on (and $\alpha_1, \alpha_2, \alpha_4, \ldots$ are treated as constants when we are integrating with respect to α_3).

After these preliminaries we are ready to explain how to construct the formula furnishing the character of $\{m - p, \mu_1, \ldots, \mu_{k-1}\}$ where $(\mu_1, \ldots, \mu_{k-1}) = (\mu)$ is a partition of p. This formula does not involve $\alpha_{p+1}, \ldots, \alpha_m$ and the part of it which is independent of $\alpha_2, \ldots, \alpha_p$ is obtained by setting $m = \alpha_1$ in Frobenius' formula for the dimension of $\{m - p, \mu_1, \ldots, \mu_{k-1}\}$. We denote the functions of $\alpha_1, \ldots, \alpha_p$ which we wish to determine by $[\mu]$. Thus $[0] = 1$, since $\{m\}$ is the identity representation, and $[1] = \alpha_1 - 1$ since the dimension of $\{m - 1, 1\}$ is $m - 1$. The part of $[2]$ which is independent of α_2 is $\frac{1}{2}\alpha_1(\alpha_1 - 3)$, since the dimension of $\{m - 2,2\}$ is $\frac{1}{2}m(m - 3)$ and so on. To obtain the part of $[\mu]$ which involves α_2 we differentiate $\{\mu\}$ with respect to 2, replace each $\{\ldots\}$ in the derivative by the corresponding $[\ldots]$ and integrate the result with respect to α_2. Thus $\{2\}_2 = \{0\} \rightarrow [0] = 1$ and so the part of $[2]$ which involves α_2 is α_2. Thus the character of $\{m - 2,2\}$ is furnished by the formula

$$[2] = \frac{1}{2}\alpha_1(\alpha_1 - 3) + \alpha_2.$$

Similarly, since $\{1^2\}_2 = -\{0\} \rightarrow -1$, the character of $\{m - 2,1^2\}$ is furnished by the formula

$$[1^2] = \frac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2) - \alpha_2.$$

For $p = 3$ we determine the part of $[\mu]$ which involves α_3 in the same way; we calculate $\{\mu\}_3$, replace each $\{\ldots\}$ in the result by the corresponding $[\ldots]$, and integrate the result with respect to α_3. If we obtain in this integral with respect to α_3 a term which already appeared in the integral with respect to α_2 we do not repeat the term. Thus the part of [3] which is independent of α_2 and α_3 is $\frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)$, since the dimension of $\{m - 3,3\}$ is $\frac{1}{6}m(m - 1)(m - 5)$; $\{3\}_2 = \{1\} \rightarrow \alpha_1 - 1$ and so the part of [3] which involves α_2 is $(\alpha_1 - 1)\alpha_2$; $\{3\}_3 = \{0\} \rightarrow 1$ and so the part of [3] which involves α_3 (and which has not already appeared in the part of [3] which involves α_2) is α_3. Thus the character of the irreducible representation $\{m - 3,3\}$ of S_m is furnished by the formula

$$[3] = \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5) + (\alpha_1 - 1)\alpha_2 + \alpha_3.$$

The formulas which we derive in this simple mechanical manner have a certain universal appeal: They are valid for all symmetric groups S_m, i.e., for every value of m. While there is no difficulty in deriving the formulas, they take up a great deal of space when p is larger than, say, 6. But this is to be expected; the symmetric group is a complex object when m is as large as, say, 16. S_{16} contains 16! elements and has 231 irreducible representations and 231 classes. Thus the mere printing of its character table is no light matter, this table being a 231×231 table involving $(231)^2$ entries. We append the formulas furnishing the characters of the irreducible representations $\{m - 5, (\mu)\}$, where (μ) is any partition of 5, the formulas for $p = 3$ and $p = 4$ being already available.[1]

$$[5] = \frac{1}{120}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 9) + \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)\alpha_2 + (\alpha_1 - 1)\binom{\alpha_2}{2} + \alpha_2\alpha_3 + \frac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_3 + (\alpha_1 - 1)\alpha_4 + \alpha_5;$$

$$[4\ 1] = \frac{1}{30}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 4)(\alpha_1 - 8) + \frac{1}{3}\alpha_1(\alpha_1 - 2)(\alpha_1 - 4)\alpha_2 - \alpha_2\alpha_3 + \frac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2)\alpha_3 - \alpha_5;$$

$$[3\ 2] = \frac{1}{24}\alpha_1(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 5)(\alpha_1 - 7) + \frac{1}{6}\alpha_1(\alpha_1 - 1)(\alpha_1 - 5)\alpha_2 + (\alpha_1 - 1)\binom{\alpha_2}{2} + \alpha_2\alpha_3 - \frac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2)\alpha_3 - (\alpha_1 - 1)\alpha_4;$$

$$[3\ 1^2] = \frac{1}{20}\alpha_1(\alpha_1 - 1)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 7) + (\alpha_1 - 1)\alpha_2 -$$

$$2(\alpha_1 - 1)\binom{\alpha_2}{2} + \alpha_5;$$

$$[2^2\,1] = \frac{1}{24}\alpha_1(\alpha_1 - 1)(\alpha_1 - 3)(\alpha_1 - 5)(\alpha_1 - 6) - \frac{1}{6}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)\alpha_2 + (\alpha_1 - 1)\binom{\alpha_2}{2} - \alpha_2\alpha_3 - \frac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_3 + (\alpha_1 - 1)\alpha_4;$$

$$[2\,1^3] = \frac{1}{30}\alpha_1(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 6) - \frac{1}{3}\alpha_1(\alpha_1 - 2)(\alpha_1 - 4)\alpha_2 + \alpha_2\alpha_3 + \frac{1}{2}\alpha_1(\alpha_1 - 3)\alpha_3 - \alpha_5;$$

$$[1^5] = \frac{1}{120}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)(\alpha_1 - 4)(\alpha_1 - 5) - \frac{1}{6}(\alpha_1 - 1)(\alpha_1 - 2)(\alpha_1 - 3)\alpha_2 + (\alpha_1 - 1)\binom{\alpha_2}{2} - \alpha_2\alpha_3 + \frac{1}{2}(\alpha_1 - 1)(\alpha_1 - 2)\alpha_3 - (\alpha_1 - 1)\alpha_4 + \alpha_5.$$

[1] Murnaghan, F. D., *Theory of Group Representations*, The Johns Hopkins Press (1938).

VECTOR FIELDS ON THE n-SPHERE

BY N. E. STEENROD AND J. H. C. WHITEHEAD

DEPARTMENT OF MATHEMATICS, PRINCETON UNIVERSITY AND OXFORD UNIVERSITY

Communicated by S. Lefschetz, November 18, 1950

1. Introduction.—Let S^n denote the n-sphere, and let 2^k be the highest power of 2 dividing $n + 1$:

$$n + 1 = 2^k(2r + 1). \tag{1}$$

Our main result is

THEOREM 1.1. *Any set of 2^k continuous vector fields tangent to S^n are somewhere dependent.*

Note first that if $r = 0$ in (1) then 1.1 is trivially true.

The case $k = 0$ of 1.1 is the classical result that a vector field on a sphere of even dimension has at least one zero. The case $k = 1$ was proved by Eckmann[1] and also by G. Whitehead.[2] A proof was given for the case $k = 2$ by G. Whitehead,[3] but was based on the erroneous assertion of Pontrjagin[4, 5] that $\pi_5(S^3) = 0$.

Whether S^n admits $2^k - 1$ independent vector fields is unsettled. It is known[1, 2] that, for $k = 1$, 2 and 3, S^n admits $2^k - 1$ such fields if $n + 1 \equiv 0 \bmod 2^k$. These fields are constructed using the complex, quaternionic and Cayley division algebras, respectively. It is not known whether there exist real division algebras on 2^k units for any $k \geqq 4$.

Before proving 1.1, we state various consequences. As shown by one of us,[6] 1.1 is equivalent to

THEOREM 1.2. *If n and k are as in (1), and if $2^k \leqq q \leqq n - 2^k$, then S^n does not admit a continuous field of tangent q-dimensional hyperplanes. Nor does S^n admit a continuous quadratic form (symmetric covariant tensor of order 2) which, at each point, is non-singular and of signature q (i.e., q negative characteristic roots).*

Let C^{n+1} be cartesian $(n+1)$-space of coordinates $(t_1, \ldots, t_{n+1})$; let S^n be the locus $\sum_1^{n+1} t_i^2 = 1$; let C^m be the subspace of the first m variables of C^{n+1} (i.e., $t_{m+1} = \ldots = t_{n+1} = 0$); let $S^{m-1} = C^m \cap S^n$; let R_{n+1} denote the rotation group (orthogonal transformations of determinant $+1$) of C^{n+1}; let R_m be the rotation group of C^m regarded as the subgroup of R_{n+1} which leaves pointwise fixed the $(n - m + 1)$-space orthogonal to C^m; and let x_i denote the vector (or point), all of whose components are zero save $t_i = 1$. Define the projection

$$p: \quad R_{n+1} \to S^n \quad \text{by} \quad p(\alpha) = \alpha(x_{n+1}), \quad \alpha \in R_{n+1}. \tag{2}$$

Then, for $x \in S^n$, $p^{-1}(x)$ is a left coset of R_n; so p is just the fibering of R_{n+1} by left cosets of R_n, and S^n is the left coset space R_{n+1}/R_n. We define the Stiefel manifold $V_{n+1,\,q+1}$ to be the left coset space

$$V_{n+1,\,q+1} = R_{n+1}/R_{n-q} \qquad (q < n). \tag{3}$$

Using inclusion relations among cosets, p decomposes into the composition of two natural projections

$$R_{n+1} \xrightarrow{g} V_{n+1,\,q+1} \xrightarrow{h} S^n. \tag{4}$$

An ordered, orthonormal set of $q+1$ vectors in C^{n+1} will be called a $(q+1)$-frame. R_{n+1} is transitive on the set of all $(q+1)$-frames based at the origin, and R_{n-q} leaves fixed the frame $(x_{n-q+1}, \ldots, x_{n+1})$. Thus $V_{n+1,\,q+1}$ can be interpreted as the manifold of $(q+1)$-frames based at the origin of C^{n+1}. The projection h assigns to each $(q+1)$-frame the end-point of its last vector. To each $(q+1)$-frame corresponds the q-frame tangent to S^n obtained by a parallel translation of the first q vectors to the end-point of the last vector. Using this correspondence, we may interpret $V_{n+1,\,q+1}$ as the manifold of all q-frames tangent to S^n, and h as the assignment of the point of tangency. By the usual process of orthogonalization, q fields of tangent vectors, which are independent at each point, can be reduced to q fields which are orthonormal at each point, i.e., to a field of tangent q-frames. Thus 1.1 is equivalent to

THEOREM 1.3. *If n and k are as in (1) with $r > 0$, then the fiber bundle $h: V_{n+1,\,2^k+1} \to S^n$ does not have a cross-section (i.e., there does not exist a continuous map $f: S^n \to V_{n+1,\,2^k+1}$ such that $hf(x) = x$ for all x).*

If f: $S^n \to R_{n+1}$ is a cross-section for $R_{n+1} \to S^n$, then gf is a cross-section for $V_{n+1,\,q+1} \to S^n$; hence

COROLLARY 1.4. *If n is not of the form $2^k - 1$, then the fiber bundle $R_{n+1} \to S^n$ does not have a cross-section.*

Using the result of Kirchhoff,[7] we have

COROLLARY 1.5. *If n is not of the form $2^k - 2$, then S^n does not admit a complex analytic structure. Even more, it does not admit an exterior form of degree 2 which is non-singular at each point.*

As shown by one of us,[8] 1.4 implies

COROLLARY 1.6. *If n is not of the form $2^k - 1$ and B is an n-sphere bundle, then the image of $\pi_n(S^n)$ in $\pi_n(B)$ under the natural map is cyclic of even or infinite order. Thus the fiber S^n is not contractible to a point in B.*

COROLLARY 1.7. *If n is not of the form $2^k - 1$, and $m > n$, then S^m cannot be fibered by n-spheres.*

COROLLARY 1.8. *A representation of S^m as an n-sphere bundle over S^{m-n} is possible only if n has the form $2^k - 1$ and $m = 2^{k+1} - 1$.*

In the last three corollaries, the group of the bundle is R_{n+1}. It is not known if they are valid with R_{n+1} replaced by the group of all homeomorphisms of S^n.

2. Squaring Operations in Projective Spaces.—Before proving 1.1, we shall establish a preliminary result. Let P^n be real projective n-space. It is known that the cohomology group $H^q(P^n)$ with coefficients mod 2 is a cyclic group of order 2 with generator $u^q (0 \leqq q \leqq n)$. The cohomology ring $H(P^n) = \sum_0^n H^q(P^n)$ satisfies $u^p \cup u^q = u^{p+q}$. For any complex K, let Sq^j: $H^q(K) \to H^{q+j}(K)$ be the squaring homomorphism determined by cup-$(q - j)$ products.[9]

THEOREM 2.1. *In P^n, we have $\mathrm{Sq}^j u^q = c(j, q)u^{q+j}$ where $0 \leqq j \leqq q$, $q + j \leqq n$, and $c(j, q)$ is the* mod 2 *value of the binomial coefficient $\binom{q}{j}$.*

The case $q = 1$ follows from the general properties:[9] $\mathrm{Sq}^0 u = u$ and $\mathrm{Sq}^q u = u \cup u$ if $q = \dim u$. Using the general formula of H. Cartan[10]

$$\mathrm{Sq}^j(u \cup v) = \sum_{i=0}^{j} (\mathrm{Sq}^i u) \cup (\mathrm{Sq}^{j-i} v),$$

and the proposition[9] $\mathrm{Sq}^i u = 0$ for $i > \dim u$, we proceed by induction:

$$\mathrm{Sq}^j u^q = \mathrm{Sq}^j(u^1 \cup u^{q-1}) = \mathrm{Sq}^0 u^1 \cup \mathrm{Sq}^j u^{q-1} + \mathrm{Sq}^1 u^1 \cup \mathrm{Sq}^{j-1} u^{q-1}$$
$$= \left\{\binom{q-1}{j} + \binom{q-1}{j-1}\right\} u^{q+j} = \binom{q}{j} u^{q+j}.$$

To apply 2.1, we shall develop a simple criterion for evaluating $c(j, q)$. It follows from 2.1 that $c(j, q)$ is the mod 2 coefficient of u^j in the expansion of $(1 + u)^q$ where $1 = u^0$ and $u = u^1$. Let

$$j = \sum_{i=0}^{m} a_i 2^i, \qquad q = \sum_{i=0}^{m} b_i 2^i$$

be the dyadic expansions of j, q (a_i, $b_i = 0$ or 1). Calculating mod 2, we have $(1 + u)^2 = 1 + u^2$; hence $(1 + u)^{2^i} = 1 + u^{2^i}$; and therefore

$$(1 + u)^q = \Pi_{i=0}^{m}(1 + b_i u^{2^i}). \tag{5}$$

By the uniqueness of the dyadic expansion, the coefficient of u^j in (5) is $b_{i_1} \cdots b_{i_p}$ where $i_1, \ldots, i_p$ are those values of i for which $a_i = 1$. Therefore

LEMMA 2.2. *$c(j, q) = 1$ if and only if the non-zero terms of the dyadic expansion of j form a subcollection of those of q.*

THEOREM 2.3. *If $n + 1 = 2^k(2r + 1)$ with $r > 0$, then, in P^n, we have*

$$\mathrm{Sq}^{2^k} u^{n-2^k} = u^n, \quad \textit{and} \quad \mathrm{Sq}^j u^{n-j} = 0 \quad \textit{for} \quad 0 < j < 2^k.$$

To prove this, note first that

$$n - 2^k = 2^{k+1}r - 1 = 2^{k+1} - 1 + 2^{k+1}(r - 1) = \textstyle\sum_{i=0}^{k} 2^i + 2^{k+1}(r - 1).$$

Therefore 2^k is a non-zero term of the dyadic expansion of $n - 2^k$. Then 2.2 implies that $c(2^k, n - 2^k) = 1$, and the first assertion follows from 2.1. To prove the second, observe that $0 < j < 2^k$ implies that j has the form $2^s(2l + 1)$ where $0 \leqq s < k$. Since

$$n = 2^k - 1 + 2^{k+1}r = \textstyle\sum_{i=0}^{k-1} 2^i + 2^{k+1}r,$$

it follows that the coefficient of 2^s in the dyadic expansion of $n - j$ is 0, while it is 1 in that of j. Then 2.2 implies $c(j, n - j) = 0$; and 2.1 applies to complete the proof.

3. *Proof of 1.1.*—We adopt the notations preceding 1.3, and assume that $r > 0$. Let E_+^m, E_-^m denote the hemispheres of S^m defined by $t_{m+1} \geqq 0$, $t_{m+1} \leqq 0$, respectively. Their totality, for $m = 0, 1, \ldots, n$, forms a cell complex on S^n.

For each $x \in S^n$, let $T(x) \in R_{n+1}$ be the composition of reflection in the n-plane through the origin perpendicular to x_1 followed by reflection in the n-plane through the origin perpendicular to x. Clearly $T(x) = T(-x)$. Since x_{n+1} is fixed under the first reflection, $pT(x)$ is the image of x_{n+1} under the second reflection. It follows that pT maps each great circle through x_{n+1} onto itself with a covering map of degree 2. From this we obtain the following properties of T:

$$T(x) = T(x') \text{ if and only if } x = x' \text{ or } x = -x', \tag{6}$$

$$T(S^n) = P^n \text{ is a real projective } n\text{-space in } R_{n+1}, \tag{7}$$

$$T(S^m) = P^m = P^n \cap R_{m+1} \text{ for } m = 0, 1, \ldots, n, \tag{8}$$

$$pT(S^{n-1}) = x_{n+1}, \text{ and } pT \text{ maps } E_+^n - S^{n-1} \tag{9}$$

topologically onto $S^n - x_{n+1}$.

Let $T' = T|S^{n-1}$. By (9), we may regard T' as a map $S^{n-1} \to R_n$. The maps T, T' were first introduced by Hopf,[11] and T' is called the characteristic or canonical map for the bundle $R_{n+1} \to S^n$. G. Whitehead[12] has shown that S^n admits q continuous, independent vector fields if and only if T' is homotopic in R_n to a map of S^{n-1} into R_{n-q}. Let

$$g': R_n \to R_n/R_{n-q} = V_{n,q}$$

be the natural map, i.e., $g' = g|R_n$. Since g' is a bundle projection, T' is deformable into a map $S^{n-1} \to R_{n-q}$ if and only if $g'T'$ is homotopic to a constant. Taking $q = 2^k$, we have only to show that $g'T': S^{n-1} \to V_{n,2^k}$ is not homotopic to a constant.

One of us[13] has given a representation of $V_{n,2^k}$ as a cell complex K consisting of one 0-cell $e^0 = g'T'(E_+^0)$, no m-cells for $0 < m < n - 2^k$, exactly one m-cell $e^m = g'T'(E_-^m)$ for $n - 2^k \leqq m \leqq n - 1$, and exactly one cell for each set of distinct integers $m_1, \ldots, m_t$ between $n - 2^k - 1$ and n, whose dimension is $\sum_1^t m_i$. The cell of least dimension greater than $n-1$ corresponds to $m_1 = n - 2^k$ and $m_2 = n - 2^k + 1$. It has dimension

$$2n - 2^{k+1} + 1 = n + 2^k(2r - 1), \qquad \text{see (1)}.$$

As $r > 0$, it follows that the n-skeleton K^n of K coincides with K^{n-1}. As $g'T'$ is a cellular map, we have $g'T'(S^{n-1}) = K^{n-1}$. Since $K^n = K^{n-1}$, we have that $g'T'$ is homotopic to a constant in $V_{n,2^k}$ if and only if it is homotopic to a constant in K^{n-1}. Thus we have only to prove that $g'T'$, regarded as a map $S^{n-1} \to K^{n-1}$, is not homotopic to a constant.

Let $e^n = gT(E_+^n) \subset V_{n+1,q+1}$, and let $K' = e^n \cup K^{n-1}$. Let $g_1: P^n \to K'$ be the map $g|P^n$. If we regard P^n as a cell complex composed of the cells $T(E_+^m)$, $m = 0, 1, \ldots, n$, then g_1 is cellular and for $n - 2^k \leqq m \leqq n$ maps the open m-cell $T(E_+^m)$ topologically onto the open m-cell e^m. It follows that g_1 induces isomorphisms of the mod 2 cohomology groups

$$g_1^*: H^i(K') \approx H^i(P^n) \qquad \text{for} \qquad n - 2^k \leqq i \leqq n. \tag{10}$$

Since $g_1^*\mathrm{Sq}^j = \mathrm{Sq}^j g_1^*$, the result of 2.3 applies with (10) to give

$$\mathrm{Sq}^{2^k} w^{n-2^k} = w^n \tag{11}$$

where w^i is the non-zero element of $H^i(K')$.

Suppose, as is impossible, that $g'T': S^{n-1} \to K^{n-1}$ is homotopic to a constant. This means that gT regarded as a map $E_+^n \to K'$ can be extended to a map $f: S^n \to K'$ such that $f(E_-^n) \subset K^{n-1}$. Clearly f^*w^n is the non-zero element of $H^n(S^n)$; and $f^*w^{n-2^k} = 0$ since $H^{n-2^k}(S^n) = 0$. But $\mathrm{Sq}^j f^* = f^*\mathrm{Sq}^j$; so we have a contradiction with (11), and the proof is complete.

REMARK 1. If, in the preceding proof, we replace 2^k by $j < 2^k$, then (11) becomes $\mathrm{Sq}^j w^{n-j} = 0$ (see 2.3), and we obtain no contradiction. This shows that the bound 2^k in 1.1 is the least value provided by our argument.

REMARK 2. Theorem 1.1, with S^n replaced by real projective space P^n, has been proved by Stiefel[14] and by Hopf.[15] Their result is a corollary of 1.1; for the double covering $S^n \to P^n$ enables one to "lift" vectors on P^n into vectors on S^n preserving independence. It may seem surprising that the theorem for P^n should be easier than for S^n. However the proof for P^n uses the fact that a certain characteristic class is non-zero. When n is odd, all the characteristic classes of S^n are zero. Thus the primary obstructions vanish, and we have shown that an obstruction of higher order is non-zero.

[1] Eckmann, B., *Comm. Math. Helv.*, **15,** 1–26 (1942).
[2] Whitehead, G., *Ann. Math.*, **43,** 132–146 (1942).
[3] Whitehead, G., *Ibid.*, **47,** 779–785 (1946); also **48,** 782 (1947).
[4] Whitehead, G., *Ibid.*, **52,** 245–247 (1950).
[5] Pontrjagin, L., *C. R. Acad. Sci. URSS,* **70,** 957–959 (1950).
[6] Steenrod, N., *Topology of Fibre Bundles,* Princeton Univ. Press, 1951, Theorems 27.16, 40.11.
[7] Kirchhoff, A., *C. R. Acad. Paris,* **223,** 1258–1260 (1948); see also reference 6, Th. 41.19.
[8] See reference 6, Th. 28.3.
[9] Steenrod, N., *Ann. Math.*, **48,** 290–320 (1947).
[10] Cartan, H., *C. R. Acad. Paris,* **230,** 425–427 (1950).
[11] Hopf, H., *Fund. Math.*, **25,** 427–440 (1935).
[12] See reference 2, Th. 10; also reference 6, Th. 27.6.
[13] Whitehead, J. H. C., *Proc. London Math. Soc.*, **48,** 243–291 (1944).
[14] Stiefel, E., *Comm. Math. Helv.*, **13,** 201–218 (1941).
[15] Hopf, H., *Ibid.*, **13,** 219–239 (1941).

MANY-FIGURE APPROXIMATIONS TO $\sqrt{2}$, AND DISTRIBUTION OF DIGITS IN $\sqrt{2}$ AND $1/\sqrt{2}$

BY HORACE S. UHLER

YALE UNIVERSITY

Communicated by J. B. Whitehead, November 15, 1950

The present investigation was begun with the single intention of testing a value for $\sqrt{2}$ as computed by J. Marcus Boorman about 65 years ago. This datum was chosen because the author had acquired the impression that Boorman seldom if ever produced a reliable result in his calculations and it seemed fair to check this particular approximation since it was claimed that he had verified many of his digits. More precisely we read[1] "Mr. J. M. Boorman of New York has computed (pp. 112–115) the square root of 2 to 520 decimals, and verified 486 of them." Then the 486 decimal digits are quoted.

At the time the author was unaware of the recent work of René Coustal[2] in which the approximation was carried to 1032 guaranteed decimal places.

This noteworthy contribution stimulated the checking of Coustal's answer and the extension of it to at least 1542 decimal places by the present critic.

The work was begun by squaring Boorman's result to find out to what extent if any the square differed from 2. This operation was conveniently effected by treating his approximation as the sum of two numbers, b and c. b represented the first 243 digits of B's $\sqrt{2}$ with the decimal point temporarily disregarded. c comprised the last 243 figures of which the first figure (8) occupied the 244th decimal place. (The 243rd place happens to be filled by zero.) The-right-hand digit of b^2 occupied the 484th decimal place while the respective leading figures of $2bc$ and c^2 fell in the 243rd and 487th places. This calculation showed that $(b + c)^2$ consisted of 1, the decimal point, and 315 nines followed by the weird sequence

59999	99987	99999	99999	99999	99999	99999	99960	00000	00000
08000	08000	00000	00000	00000	00000	00000	00000	00000	00000
00000	00069	99999	99998	80000	00000	00000	59999	99999	99999
99999	99999	99999	99996	43244	95623	60074	24492	20770	18867
92230	99222	52571	61930	68368	12237	83658	55185	33491	46724
64379	19306	83618	72765	59353	71418	55794	54954	20347	·····

Hence it has now been proved that Boorman's *revised* value of $\sqrt{2}$ was correct to 315 decimals but erroneous for the remaining 171 figures.

The next stage of the present work comprised the following steps, the theory and justification for which will be outlined below. The preceding value of $(b + c)^2$ was subtracted from 2, this difference was divided by 4, and then $[2 - (b + c)^2]/4$ was multiplied by the 315-place correct approximation to $\sqrt{2}$. In this way the author obtained the value of $\sqrt{2}$ to 611 decimal places.

About the same time Dr. John W. Wrench, Jr., computed an approximation to $\sqrt{2}$. On August 27, 1950, he juxtaposed his value against the 611-place work strip and found complete agreement as far as his last digit in the 460th decimal place.

Let a denote a definite approximation to $\sqrt{N}$ and ϵ the correction to a to give a higher approximation. More precisely

$$\sqrt{N} = a + \epsilon. \tag{1}$$

(Expressed in decimals $\sqrt{N}$ and ϵ would never terminate when N is not a perfect square.) In practice we square a and subtract the result from N. Let this difference be symbolized by δ so that

$$N - a^2 = \delta. \tag{2}$$

Now $\epsilon = \sqrt{N} - a$ from equation (1) and $\sqrt{N} \equiv a/\sqrt{a^2/N} = a(1 - \delta/N)^{-1/2}$ from equation (2). The condition for convergence $\delta^2/N^2 < 1$ should be amply fulfilled so that the required practical expansion may be written

$$\epsilon = \frac{a\delta}{2N}\left\{1 + \frac{3\delta}{4N} + \frac{5\delta^2}{8N^2} + \frac{35\delta^3}{64N^3} + \cdots\right\}. \qquad (3)$$

In the present case $N = 2$ and the first order term applied earlier is $\epsilon = a(\delta/4)$.

The author's 612 digit approximation was identical with the correspond-

TABLE 1

$\sqrt{2}$ =

1.41421	35623	73095	04880	16887	24209	69807	85696	71875	37694
80731	76679	73799	07324	78462	10703	88503	87534	32764	15727
35013	84623	09122	97024	92483	60558	50737	21264	41214	97099
93583	14132	22665	92750	55927	55799	95050	11527	82060	57147
01095	59971	60597	02745	34596	86201	47285	17418	64088	91986
09552	32923	04843	08714	32145	08397	62603	62799	52514	07989
		B←							
68725	33965	46331	80882	96406	20615	25835	23950	54745	75028
77599	61729	83557	52203	37531	85701	13543	74603	40849	88471
60386	89997	06990	04815	03054	40277	90316	45424	78230	68492
	W←								
93691	86215	80578	46311	15966	68713	01301	56185	68987	23723
52885	09264	86124	94977	15421	83342	04285	68606	01468	24720
77143	58548	74155	65706	96776	53720	22648	54470	15858	80162
07584	74922	65722	60020	85584	46652	14583	98893	94437	09265
91800	31138	82464	68157	08263	01005	94858	70400	31864	80342
19489	72782	90641	04507	26368	81313	73985	52561	17322	04024
50912	27700	22694	11275	73627	28049	57381	08967	50401	83698
68368	45072	57993	64729	06076	29969	41380	47565	48237	28997
18032	68024	74420	62926	91248	59052	18100	44598	42150	59112
02494	41341	72853	14781	05803	60337	10773	09182	86931	47101
71111	68391	65817	26889	41975	87165	82152	12822	95184	88472
						C←			
08969	46338	62891	56288	27659	52635	14054	22676	53239	69461
75112	91602	40871	55101	35150	45538	12875	60052	63146	80171
27402	65396	94702	40300	51749	53188	62925	63138	51881	63478
00156	93691	76881	85237	86840	52287	83762	93892	14300	65586
95686	85964	59515	55016	44724	50983	68960	36887	32311	43894
15576	65104	08839	14292	33811	32060	52433	62948	53170	49915
77175	62285	49741	43899	91880	21762	43096	52065	64211	82731
67262	57539	59471	72559	34637	23863	22614	82742	62220	86711
55839	59992	65211	76252	69891	75409	88159	34864	00834	57085
18147	22318	14204	07042	65090	56532	33339	84364	57865	79679
65192	67292	39987	53666	17215	98257	88602	63363	6178	

ing portion of Coustal's longer value. This agreement justified the application of formulae (2) and (3) to the 612 figure value of a. The second calculation was carried to the 1040th decimal place. The result previously obtained by Coustal agreed absolutely with this slightly more extended datum to 1034 places. This circumstance fully confirms Coustal's

claim that "Tous ces chiffres sont exacts, la 1033^{e} décimale est sûrement un zéro, la 1034^{e} est probablement un 5." Hence "probablement" may now be changed to "certainement."

In this case $\delta = (0.161859\cdots) \times 10^{-610}$; $a\delta/4 = (0.572260\cdots) \times 10^{-611}$, i.e., 611 decimally antecedent zeros; $3a\delta^2/32 = (3/32)(1.414213\cdots)(1.618\cdots)^2 \times 10^{-1222} = (0.347346\cdots) \times 10^{-1222}$, i.e., 1222 ante-zeros which safely include the 1034th decimal place.

Finally in order to contribute something entirely new, while using a machine in the same non-electronic class as that presumably employed by Coustal, the perfect 1033 digits of $\sqrt{2}$ were assigned to a. Then the 1544 decimal approximation to $\sqrt{2}$ as given below was derived from

TABLE 2

n	x_n^2	p_n	x_n^2	p_n
100	7.6	0.4240	5.4	0.2068
200	5.1	0.1804	3.4	0.0544
300	5.46	0.2125	1.8	0.0059
400	7.2	0.3831	3.25	0.0470
500	3.92	0.0853	1.56	0.0039
600	4.9	0.1634	2.4	0.0176
700	6.914···	0.3539	3.371···	0.0527
800	7.075	0.3703	2.025	0.0091
900	8.28	0.4945	3.24	0.0468
1000	8.18	0.4833	3.06	0.0394
1100	8.581	0.5196	3.90	0.0846
1200	9.316	0.5799	3.883	0.0831
1300	8.323···	0.4980	5.615···	0.2254
1400	8.957···	0.5504	4.157···	0.0993
1500	9.12	0.5638	3.093	0.0407
1545	9.401···	0.5869	1.776···	0.0056

equations (2) and (3). The superior letters B, W, and C in table 1 mark the limits of accuracy attained, respectively, by Boorman, Wrench and Coustal.

In addition to checking throughout all stages of the work with small auxiliary moduli of the form $10^n + 1$ the final answer was carefully squared. This long multiplication gave an unbroken succession of 1544 nines between the decimal point (after 1) and a 2. The writer guarantees 1542 decimal figures and he believes that the digits in the 1543rd and 1544th places are correct.

In the last case $\delta = (0.153376\cdots) \times 10^{-1032}$; $a\delta/4 = (0.542267\cdots) \times 10^{-1033}$, i.e., 1033 decimally antecedent zeros; $3a\delta^2/32 = (3/32)(1.414213\cdots)(1.533764\cdots)^2 \times 10^{-2066} = (0.311891\cdots) \times 10^{-2066}$, i.e., 2066 ante-zeros which overwhelm the 1544th decimal place.

Table 2, given above, relative to the distribution of the digits of $\sqrt{2}$ and $1/\sqrt{2}$ may possess statistical interest. The values of p_n were obtained by linear interpolation in a table of percentage points of the χ^2 distribution by Catherine M. Thompson.[3] However the value of P for $n = 1$ and nine degrees of freedom could not be located by the writer, hence to avoid extrapolation he calculated the product

$$p = \frac{1}{2^{4.5}\Gamma(4.5)} \int_0^1 e^{-x/2} x^{3.5}\, dx,$$

with the aid of a thirteen-place table of logarithms by H. Andoyer (1922), and found $p = 0.000562497302$.

In table 2 the second and third columns pertain to $\sqrt{2}$ and they indicate settling down to a fairly steady normal state for $n > 800$. The fourth and fifth columns refer to $1/\sqrt{2}$ and, with two or three exceptions, they show noteworthy small values of χ_n^2 and p_n. The relatively large values of p_n for $n = 1100$, 1200 and 1300 can be traced back to very appreciable steady increases in the number of 4's combined with simultaneous deficiencies in the number of 9's occurring.

[1] Editor, *Mathematical Magazine*, **1**, 164 (1887).
[2] Coustal, René, *Compt. rend.*, **230**, 431, 432 (1950).
[3] Thompson, Catherine M., *Biometrika*, **32**, 187–191 (1941–1942).

INFORMATION TO CONTRIBUTORS

THE PROCEEDINGS is the official organ of the NATIONAL ACADEMY OF SCIENCES and of the NATIONAL RESEARCH COUNCIL for the publication of brief accounts of important current researches of members of the ACADEMY and of the COUNCIL and of other American investigators. The PROCEEDINGS will aim especially to secure prompt publication of original announcements of discoveries and wide circulation of the results of American research among investigators in other countries and in all branches of science.

ARTICLES should be brief. The viewpoint should be comprehensive in giving the relation of the paper to previous publications of the author or of others and in exhibiting, where practicable, the significance of the work for other branches of science. Elaborate technical details of the work and long tables of data should be avoided, but authors should be precise in making clear the new results and should give some record of the methods and data upon which they are based.

MANUSCRIPTS should be prepared with a current number of the PROCEEDINGS as a model in matters of form, and should be typewritten in duplicate with double spacing, the author retaining one copy. Illustrations should be confined to text figures of simple character, though more elaborate illustration may be allowed in special instances to authors willing to pay for their preparation and insertion. Particular attention should be given to arranging tabular matter in a simple and concise manner.

REFERENCES to literature, numbered consecutively, will be placed at the end of the article and short footnotes should be avoided. It is suggested that references to periodicals be furnished in some detail and in general in accordance with the standard adopted for the Subject Catalogue of the International Catalog of Scientific Literature, viz., name of author, with initials following (ordinarily omitting title of paper), abbreviated name of Journal, with place of publication, series (if any), volume, inclusive pages, year. For example: Montgomery, T. H., *J. Morph.*, *Boston*, 22, 731–815 (1911); or, Wheeler, W. M., *Konigsberg, Schr. physik. Ges.*, 20, 1–142 (1914).

MANUSCRIPTS may be sent to the Managing Editor, Edwin Bidwell Wilson, Harvard School of Public Health, 695 Huntington Ave., Boston 15, Mass., only by members of the ACADEMY or by chairmen of the divisions of the NATIONAL RESEARCH COUNCIL who will assume responsibility for the propriety and scientific standards of the paper. Articles are limited to six printed pages; *excess space is billed to the author at $14.00 per page.*

PROOF will ordinarily be sent,—with the understanding that charges for an author's corrections shall be billed to him. Authors are therefore requested to make final revisions on the typewritten manuscripts.

REPRINTS should be ordered at the time of submission of manuscript. They will be furnished to authors at cost, approximately as follows:

SCHEDULE OF RATES FOR REPRINTS, CARRIAGE CHARGES EXTRA

No. Copies	50	100	150	200	250
4 pp.	4.18	4.72	5.30	5.90	6.50
8 pp.	8.26	9.44	10.62	11.79	12.97
Covers	3.15	4.03	4.92	5.81	6.69

CONTENTS

		Page
ASTRONOMY.—ON THE ORIGIN OF THE SOLAR SYSTEM	*By Gerard P. Kuiper*	1
ASTRONOMY.—PROXIMA CENTAURI AS A FLARE STAR	*By Harlow Shapley*	15
ASTRONOMY.—THE THEORY OF MICRO-METEORITES. PART II. IN HETEROTHERMAL ATMOSPHERES	*By Fred L. Whipple*	19
BIOCHEMISTRY.—MULTIMOLECULAR ADSORPTION OF HORSE HEART METMYOGLOBIN ONTO FILMS OF BARIUM STEARATE	*By Albert A. Fish*	30
BIOCHEMISTRY.—A COMPARISON OF THE CONTENT OF DESOXYRIBOSENUCLEIC ACID (DNA) IN ISOLATED ANIMAL NUCLEI BY CYTOCHEMICAL AND CHEMICAL METHODS	*By Cecilia Lenchlenberger, Roger Vendrely and Colette Vendrely*	33
GENETICS.—STRUCTURES IN E. COLI RESEMBLING CHROMONEMATA	*By Alfred Marshak*	38
MATHEMATICS.—PARTIAL DIFFERENTIAL EQUATIONS AND GENERALIZED ANALYTIC FUNCTIONS. SECOND NOTE	*By Lipman Bers*	42
MATHEMATICS.—THE HARMONIC OPERATOR FOR EXTERIOR DIFFERENTIAL FORMS	*By Matthew P. Gaffney*	48
MATHEMATICS.—ON THE ANALYSIS OF REPRESENTATIONS OF THE LINEAR GROUP	*By Francis D. Murnaghan*	51
MATHEMATICS.—THE CHARACTERS OF THE SYMMETRIC GROUP	*By Francis D. Murnaghan*	55
MATHEMATICS.—VECTOR FIELDS ON THE n-SPHERE	*By N. E. Steenrod and J. H. C. Whitehead*	58
MATHEMATICS.—MANY-FIGURE APPROXIMATIONS TO $\sqrt{2}$, AND DISTRIBUTION OF DIGITS IN $\sqrt{2}$ AND $1/\sqrt{2}$	*By Horace S. Uhler*	63